

CHARACTER AS NARRATIVE:  
MORAL RESPONSIBILITY IN CONTEXT

Elizabeth Cargile Williams

Submitted to the faculty of the Graduate School in partial  
fulfillment of the requirements  
for the degree  
Doctor of Philosophy  
in the Department of Philosophy,  
Indiana University  
May 2024

Accepted by the faculty of the University Graduate School, Indiana University, in partial fulfillment of the requirements for the degree Doctor of Philosophy.

**Doctoral Committee:**

Kate Abramson, Ph.D., Chair (Philosophy) \_\_\_\_\_

Gary Ebbs, Ph.D. (Philosophy) \_\_\_\_\_

Adam Leite, Ph.D. (Philosophy) \_\_\_\_\_

Cynthia Wu, Ph.D. (Gender Studies) \_\_\_\_\_

**Defense Date:** April 23, 2024

Copyright © 2024  
Elizabeth Cargile Williams

## **Acknowledgements**

Thank you to my advisor, Kate Abramson, for being patient with me throughout this long and winding process. All your feedback was constructive, and I always knew that you were there to help with whatever battle I was fighting at the time.

Thank you to my second reader Gary Ebbs for challenging philosophical comments that developed this project significantly. Thank you to Adam Leite for insightful remarks on early drafts. And finally, thank you to Cynthia Wu for showing me what greater freedom in writing style could look like in an academic setting.

This project would not have come to completion without the encouragement and philosophical advice of fellow graduate students Ricky Mouser, Savannah Pearlman, and Shannon Abelson.

I would not have walked away from grad school debt free without the ongoing financial and moral support of my parents and grandparents.

A special thank you to my therapists, my good friend Jordan Baker, and my dog Kensie for their ongoing emotional support throughout this four-year project.

Elizabeth Cargile Williams

CHARACTER AS NARRATIVE: MORAL RESPONSIBILITY IN CONTEXT

In this dissertation, I propose a theory of moral responsibility for character, understood as a kind of narrative. Narratives of character give evaluative and emotional coherence to the judgments and choices of others understood within context. In Chapter 1, I set the stage for the project by arguing that Gary Watson's (1996) notion of accountability, properly understood, requires some reference to the particular person, rather than merely roles or rules. In Chapter 2, I push back against R. Jay Wallace's (1994) view of responsibility that makes us out to be responsible only for our choices or the things we can directly control. In Chapter 3, I argue that T. M. Scanlon's (1998) view rightly captures responsibility for judgment-sensitive attitudes but wrongly underplays the distinct importance of choice and control. In Chapter 4, I consider Susan Wolf's (1990) and Fischer and Ravizza's (1998) historical requirements on responsibility and conclude that the former too quickly removes agency and responsibility from those who have chosen to go down evil paths while the latter isn't something we have access to in our everyday moral practices of accountability. In Chapter 5, I draw on Peter Goldie's (2012) work on narratives to create a novel view of character, understood not as a stable disposition but as an actively constructed narrative built from a number of elements we care about, including our choices, our judgments, our history and context, as well as other features of what we're like. These narratives often draw out patterns that have interpersonal import, and they can range from very small character studies to narratives that span a person's life. Finally, in Chapter 6, I argue that constructing a taxonomy that clearly delineates who is responsible from who is not is a project that is doomed from the start, and I consider what my broadly Strawsonian approach can tell us about the excuses and exemptions.

Acceptance Page.....	ivi
Copyright Page .....	ivii
Acknowledgements .....	iv
Abstract.....	iv
Table of Contents.....	ivi
Introduction.....	1
Chapter 1: Attributability, Accountability, and the Importance of Character.....	5
Part 1: Attributability and Accountability .....	6
1.1 The Watsonian Inspiration .....	6
1.2 Zheng’s Characterization of Accountability.....	8
1.3 Why We Should Separate Accountability and Substantive Responsibility .....	10
1.4 Moral Accountability Requires Moral Attributability .....	14
1.5 Delimiting the Question .....	15
Part 2: Responsibility in Context .....	15
2.1 A Range of Starting Questions.....	16
2.2 Common Pitfalls.....	16
2.3 Why Does Responsibility Matter?.....	18
2.4 The Core of Moral Responsibility.....	20
Conclusion.....	22
Chapter 2: Wallace’s Choice-Centered Account of Moral Responsibility .....	23
Part 1: Wallace.....	23
1.1 Wallace’s Larger Project.....	23
1.2 Wallace’s First Argument.....	24
1.3 Wallace’s Second Argument .....	26
1.4 Summary .....	29
Part 2: Departures from Wallace .....	31
2.1 Departures from Wallace’s First Argument .....	31
2.2 Departures from Wallace’s Second Argument.....	36
2.3 Conflicts Between Attitudes and Choices .....	37
2.4 Tracing Accounts .....	41
2.5 Worries About Moralism .....	43
Conclusion.....	45
Chapter 3: Scanlon’s Judgment-Centered Account of Moral Responsibility .....	46
Part 1: Scanlon.....	46
1.1 Substantive Responsibility and Attributability Responsibility .....	46

1.2 Scanlon’s Judgment-Centered View of Moral Responsibility.....	48
1.3 Scanlon vs. Wallace .....	52
Part 2: Departures from Scanlon .....	53
2.1 Direct and Indirect Control.....	54
2.2 Varieties of Indirect Control .....	56
2.3 Embodiment.....	57
2.4 Historical Context .....	58
2.5 Inferring Judgments .....	60
Conclusion.....	62
Chapter 4: How History and Context Matter .....	64
Part 1: Wolf’s Reason View.....	64
1.1 Wolf’s Reason View .....	64
1.2 Mental Illness.....	66
1.3 Poor Upbringing.....	68
1.4 Hardening Your Heart.....	71
Part 2: Fischer and Ravizza’s Historical View .....	73
2.1 Fischer and Ravizza’s Historical Account .....	74
2.2 Taking Responsibility.....	77
2.3 Lessons for a Positive View.....	79
Conclusion.....	80
Chapter 5: Responsibility for Character .....	82
Part 1: Narratives of Character .....	83
1.1 A Few Preliminaries .....	83
1.2 Common Narratives.....	86
1.3 The Details We Care About .....	89
1.4 Getting the Interpretation(s) Right.....	94
Part 2: Interesting Upshots .....	98
2.1 Incomplete, Unstable Character.....	98
2.2 Activity and Passivity in Moral Life .....	100
2.3 Character as Narrative; Not the Essence that Underlies the Narrative .....	102
2.4 Relativism? and Narratives of Character .....	104
Conclusion.....	105
Chapter 6: Excuses and Exemptions? .....	106
Part 1: Why a Taxonomy Isn’t a Good Idea.....	106
1.1 Character as a Mode.....	107

1.2 Carving off the Bounds of Responsibility .....	110
Part 2: Strawson and the Limits of Responsibility .....	115
2.1 OCD and Responsibility.....	115
2.2 Excuses.....	117
2.3 Exemptions.....	120
2.4 Ameliorating Strawson? .....	122
Conclusion.....	125
Conclusion.....	126
References.....	129
Curriculum Vitae.....	



## Introduction

“Be a philosopher; but, amidst all your philosophy, be still a man.”

~David Hume<sup>1</sup>

When I started this project, I wanted to determine how to paint the most exacting and complete picture of character so that I could properly blame myself and determine areas for improvement. I don't know that I've freed myself from this perfectionist line of thought, but I've come to see my task as contextualizing responsibility rather than precisifying it at a microscopic scale. In developing my view, I've tried to articulate to myself why character matters to us in our interpersonal relationships, what things we care about when it comes to character, and when adopting the mode of character is appropriate.

In this dissertation, I build, step-by-step, an alternative theory of moral responsibility for character, understood as a kind of narrative. Narratives of character give evaluative and emotional coherence to the judgments and choices of others understood within context. In the case of moral responsibility, that evaluative and emotional coherence invokes the reactive attitudes. I take R. Jay Wallace's (1994) and T.M. Scanlon's (1998) accounts of moral responsibility for choices and judgments, respectively, as my starting points. I build on their work and the work of others to try to capture the full range of things we care about in the narratives of character we tell about ourselves and each other. My account is explicitly anti-reductionist and leaves significant room for the application of practical judgment.

Throughout the dissertation, I speak of our moral practices. Whatever view I come to will be revisionary of at least some part of what we take to be our ordinary moral practices, which are often

---

<sup>1</sup> *EHU 1.6/SBN 9*

divided and conflicted in interesting ways. This “we” is more aspirational than descriptive. The hope is that the view I am describing and the cases I am considering will be appealing as we come to fully appreciate them. The real test, however, is whether living out the philosophical commitments in this view makes our lives better. For my own part, it has been a marked improvement over my previously held views.

In Chapter 1, I set the stage for the project by intervening in the literature on Gary Watson’s (1996) distinction between accountability and attributability, arguing that moral accountability, properly understood, requires some reference to the particular person, rather than merely roles or rules. I also discuss a range of common methodological approaches to thinking about moral responsibility that make responsibility out to be a burden by focusing so intently on the fairness of blame and sanctions. Instead, I present a basically Strawsonian frame for moral responsibility that places gratitude and resentment as the paradigmatic pair of reactive attitudes that carve out the center of our moral practices, though it is inclusive of a wide range of attitudes around and between.

In Chapter 2, I consider R. Jay Wallace’s view of responsibility that makes us out to be responsible only for our choices or the things we can directly control. I argue that we can also be responsible for our judgments and attitudes that are not always under our direct control. I consider whether tracing arguments can be employed to preserve the centrality of choice and conclude that choice and direct control are not necessary, even in the background, for someone to be responsible for their attitudes and judgments. Finally, I discuss worries about moralism that arise when we allow for the possibility of holding others responsible for attitudes that aren’t under their direct control and that they are actively working against. On an interpersonally-focused view, I show that we would need good reason, on balance, to consider these kinds of flaws and hold others responsible for them. Otherwise, making those kinds of judgments can quickly become a vice of something like busybodiness, excessive self-righteousness, or over-scrupulousness.

In Chapter 3, I turn to T. M. Scanlon's view of responsibility for judgment-sensitive attitudes. While Scanlon captures responsibility for those things that are under our indirect control, his view underplays the distinct importance of choice and control. I concur with Robert Adams (1985) that both choice and judgment are necessary to understand the complicated internal workings of self-governance and that neither should be given priority over the other. Scanlon's reasons-based picture of responsibility also appears to be disconnected from who we are as embodied creatures with historical contexts and sometimes opaque internal lives. I go beyond Scanlon to look at responsibility for character as something that is deeply dependent on things like our histories and who we take ourselves to be.

In Chapter 4, I consider Susan Wolf's (1990) and Fischer and Ravizza's (1998) historical views of responsibility. I reject Wolf's view that we must be able to understand what is true and right and in accordance with Reason, as it too quickly removes agency and responsibility from those who have chosen to go down evil paths. I consider three kinds of cases that illustrate the limits of Wolf's view, including mental illness, a poor upbringing, and those who harden their hearts. I reject Fischer and Ravizza's historical condition that requires an agent to have taken responsibility at some point, as it isn't something we have access to in our everyday moral practices of accountability. The background dynamics of who is expected to take responsibility are often complex and dependent on power and gender norms in a way that further complicates Fischer and Ravizza's picture.

In Chapter 5, I draw on Peter Goldie's (2012) work on narratives to create a novel view of character, understood not as a stable disposition but as an actively constructed narrative built from a number of elements we care about, including our choices, our judgments, our history and context, as well as other features of what we're like. These narratives often draw out patterns that have interpersonal import, and they can range from very small character studies to narratives that span a person's life. My view has a number of interesting upshots, including the conclusions that character

can be incomplete and unstable and that our narratives of character are always open-ended. While others have tried to defend responsibility for judgments and attitudes as a kind of activity that is up to us in some important sense, I conclude that responsibility for character is something that we are both active and passive with respect to. I also defend the theoretical separation of character as a narrative from the lives we actually live and conclude that we do not need to refer to some deeper essence to justify the narratives of character we tell.

Finally, in Chapter 6, I argue that constructing a taxonomy that clearly delineates who is responsible from who is not based on medical categories is a project that is doomed from the start, and I consider what my broadly Strawsonian approach can tell us about the excuses and exemptions. I borrow heavily from the work of Kate Abramson (2016) to argue that character is a mode of evaluation and that choosing which mode of evaluation to apply (including the medical and natural ability modes, among others) is itself fraught with ethical questions all the way down. I also consider David Shoemaker's (2015) approach to thinking about responsibility at the margins and conclude that such a project is not advisable nor feasible, as taxonomies like Shoemaker's run afoul of a host of philosophical and practical issues. I conclude by using the case of an agent with moral scrupulosity OCD to demonstrate how we should interpret the Strawsonian frame and apply it according to ethical considerations rather than medical categories.

## Chapter 1: Attributability, Accountability, and the Importance of Character

How central are blame and sanctions to our practices of moral responsibility? And how does character fit into the equation?

Gary Watson's (1996) paper "Two Faces of Responsibility" distinguishes moral attributability from moral accountability. In Watson's terminology, responsibility as attributability concerns questions about how an agent's actions reflect upon her character or practical identity. These assessments warrant judgments of fault and/or blame without a further tendency towards sanctions or adverse treatment. Accountability, on the other hand, involves sanctions. On this face of responsibility, sanctions are only fair and appropriate if the agent is able to avoid them.

In this chapter, I aim to reclaim the attributability face of responsibility contra other contemporary accounts that advocate for moral accountability without attributability. First, I show that something crucial about our moral practices is lost when we jettison attributability. Second, I demonstrate that focusing on the fairness of blame and sanctions alone makes responsibility out to be a burden.

This chapter sets the stage for a broadly virtue theoretical project, in which I attempt to synthesize recent views on moral responsibility to create a compelling account of character, understood as a type of narrative that typically includes patterns of our choices and judgments understood in context.<sup>2</sup> By reconceptualizing the connection between attributability and accountability, I motivate why judgments of and responses to character go beyond mere grading against external standards of excellence. In the end, our practical identities are deeply bound up in our practices of moral responsibility, which support and perhaps may even make possible certain kinds of interpersonal relationships.

---

<sup>2</sup> Note: Context does not always excuse. Sometimes it reveals greater blameworthiness. More on this in Chapter 4.

## Part 1: Attributability and Accountability

In Part 1 of this chapter, I demonstrate that holding a particular person accountable (moral accountability) is relevantly different from creating a set of burdens and obligations that apply to different roles (substantive responsibility). While Watson's (1996) and Scanlon's (1998) notions of accountability and substantive responsibility appear to overlap within the domain of moral responsibility, I claim that moral accountability requires a specific kind of interpersonal address that is lacking in substantive responsibility. These two kinds of responsibility may overlap, but they are distinct.

Substantive responsibility does not require any assessment of character or practical identity; moral accountability does. As such, moral accountability requires attributability. This revised notion of moral accountability points us in a virtue theoretic direction, re-centering character and re-emphasizing the kinds of interactions that happen between concrete, instantiated persons.

I take my inspiration from Gary Watson's (1996) claims that accountability blame requires attributability blame and that moral accountability is not exhausted by questions about the appropriateness of sanctions. My view pushes back against Robin Zheng's (2016) characterization of moral accountability which is spelled out roughly as a version of T. M. Scanlon's (1998) notion of substantive responsibility. On Zheng's understanding of accountability, accountability and attributability come apart.

### 1.1 The Watsonian Inspiration

Watson (1996/2004) identifies two faces of responsibility: attributability and accountability. On Watson's conception, responsibility as attributability is a kind of aretaic appraisal—when we judge someone responsible in the attributability sense, we consider how her actions reflect upon her

character, her practical identity, or the ends she has adopted.<sup>3</sup> Attributability (or aretaic) blame and/or judgments of fault may be appropriate for this face of responsibility, but sanctions can only enter in at the level of accountability.<sup>4</sup> This is because sanctions raise questions about fairness, which requires avoidability. Accountability, then, requires some kind of control that the agent can exercise to avoid placing themselves in a position in which they might encounter sanctions.<sup>5</sup> Moral sanctions can include accountability blame or other adverse treatment (276).<sup>6</sup> These two kinds of blame are not equivalent, as aretaic blame (or attributability blame) is distinguished from accountability blame in that accountability blame only comes into play when “some *further response* to the agent is (in principle) appropriate” (265).

Both notions of responsibility can be understood across domains. However, when it comes to the moral domain, Watson suggests that accountability has important features that may not always be captured by non-moral kinds of accountability. Take the example of a criminal who was raised in squalid circumstances and who has been unfairly treated at every turn. In this case, we might think that accountability blame is inhibited by the thought that the criminal deserves pity rather than blame. At the same time, it seems appropriate to inflict legal sanctions.

---

<sup>3</sup> Note that Watson’s view of the aretaic is broader than most discussions of the aretaic and does not only concern virtues or vices that an action might reveal (courage, sloth, etc.) but also an agent’s practical identity and the ends she might adopt. For more discussion of aretaic appraisal, see Mason (2006).

<sup>4</sup> Angela Smith (2008) worries that Watson’s view of attributability amounts to mere grading, e.g., x is taller than y, not the deep responsibility that is moral responsibility: “the notion of aretaic appraisal would seem to be very close to Smart’s idea of praise and dispraise. Aretaic appraisals would report how a person stands with regard to certain standards of human excellence, but would not bring with them any further implications of fault or discredit, and would not support any demands for reasonable regard on the part of others. But if aretaic evaluations are simply ‘grading’ evaluation, which do not imply any normative failure on the part of the person appraised, then Watson’s insistence that they are nevertheless ‘deep’ forms of assessment becomes considerably more dubious” (377-378). For more on praise and dispraise, see J. J. C. Smart (1961).

<sup>5</sup> According to Watson, “*It is unfair to impose sanctions upon people unless they have a reasonable opportunity to avoid incurring them.* If one had a fair opportunity to avoid being subject to the demand in the first place, or if, once subjected, one has the capacity to comply with it, then this condition is met. If one is subject to a requirement as part of a noncoercive and nonexploitative agreement, then one might be fairly held responsible for failing to satisfy it even if ... one is quite unable to do so” (276).

<sup>6</sup> For accounts that press back against Watson’s conception in different ways, see Smith (2008; 2015a) and Shoemaker (2011). Instead of Watson’s two faces of responsibility, Smith favors the singular face of answerability. Shoemaker, on the other hand, favors three faces of responsibility.

Once we jettison accountability blame, our sanctioning responses “seem regulative rather than retributive. In a disconcerting way, they lose their normal expressive function” (281). Focusing on sanctions alone treats “moral accountability as a legal-like practice,” which “leaves out crucial features of moral blame” (280). Though accountability can concern sanctions or blame, there is a distinct relation between accountability and attributability when it comes to blame: “accountability blame is a response to the faults identified in aretaic blame” (278).

I agree with Watson on at least three things: First, there is an important expressive function of sanctions that should be preserved in the moral case. Second, focusing solely on sanctions in the moral case seems to problematically reduce moral accountability to a legal-like practice. Third, there is an important relation between accountability and attributability that we should try to preserve.

One small note before I turn to Zheng: It is not my project here to strongly separate Watson’s view from Zheng’s. Likewise, I will set aside questions as to how closely Watson’s notion of accountability lines up with Scanlon’s notion of substantive responsibility.

## 1.2 Zheng’s Characterization of Accountability

Zheng’s (2016) view interprets moral accountability largely in terms of Scanlon’s notion of substantive responsibility.<sup>7</sup> According to Zheng, “[w]e are morally responsible for an action in this accountability sense when it is appropriate for others to hold us to certain expectations and demands regarding our duties and tasks—and to sanction us when we fail to carry them out” (66). What matters on this conception is that one has taken on certain role responsibilities that carry with them

---

<sup>7</sup> Zheng states that Scanlon’s (1998) notion of substantive responsibility “roughly corresponds to my concept of accountability” (74). We can see the connection if we look at Scanlon’s characterization of substantive responsibility as concerning “substantive claims about what people are required (or, in this case, not required) to do for each other” (248). Scanlon’s conception primarily concerns the burdens and obligations that we would not reasonably reject.



certain burdens (63). None of this requires an assessment of character or blameworthiness. As such, it does not require attributability.

Zheng's framing of accountability in terms of substantive responsibility (and subsequent separation of accountability from attributability) is motivated by cases of implicit bias. It is often unclear whether implicit biases are attributable to agents, and yet those implicit biases can result in significant negative consequences for others. Substantive responsibility allows us to hold people responsible for the consequences of their implicit biases by asking them to change their practices, make reparations, or do something else to repair the harm done. Likewise, moving past questions of blame can perhaps allow us to find more productive ways to interact with those who exhibit implicit bias and who may be resistant to accepting moral criticism for their behavior.

Zheng illustrates her view with Scanlon's discussion of the milk-seller, who, despite taking all reasonable precautions, has sold contaminated milk:

On the strict liability model, she remains accountable because she had the opportunity to avoid incurring such burdens, i.e., she could have refrained from entering the milk business. Entering the milk business, in effect, is entering what Scanlon calls an 'affected area': those who do so lay down their rights not to be penalized for violating the regulations that govern the affected area (74).

Since these penalties do not require moral fault, substantive responsibility in the absence of attributability "requires strong justification; the social goals (e.g., public safety) it promotes must be sufficiently important to outweigh the risks imposed on people who enter the affected area" (74).

There are some reasons that suggest there is daylight between Zheng's and Scanlon's views, though the two are closely related. For instance, Zheng places more emphasis on role responsibilities than Scanlon. Scanlon also makes several remarks that suggest he shares Watson's worries about cases like strict liability. First, Scanlon states that legal blame is only appropriate if something wrong has been done. If the milk-seller's punishment were an expression of "legal blame," then sanctions would not be justified (266). Second, Scanlon worries more generally that the

“enlargement of the affected area to include morally unobjectionable activities” makes the laws or other principles that might recommend sanctions for engaging in those morally unobjectionable activities harder to justify (266). Finally, when it comes to moral blame, Scanlon suggests that it should be understood as “fundamentally a judgment of condemnation, not a penalty” (267).

I’m sympathetic to Zheng’s idea that we should expand our forms of interpersonal address beyond a repertoire of blame, but I think there are good reasons to keep moral accountability separate from substantive responsibility. Before I provide these reasons, it is important to note that accountability, on Zheng’s view, is interpersonal in an important sense. For Zheng, accountability is “primarily a matter of interpersonal, not metaphysical, relations; not what it takes to be an agent, but what it takes to be a member of a community of agents” (67). This comports with Scanlon’s notion of substantive responsibility, which depends on agreement upon a set of principles that no one could reasonably reject, or what we owe to each other. However, as I will argue, the interpretation of accountability in terms of substantive responsibility is not interpersonal in the right way.

### 1.3 Why We Should Separate Accountability and Substantive Responsibility

None of what I am about to say is necessarily a critique of substantive responsibility. My aim is to call attention to an important set of differences between our normal practices of moral accountability and the allocation of burdens based on role responsibilities. These differences suggest we should separate accountability from substantive responsibility within the moral domain.

If we understand accountability in the way Zheng suggests, we end up with the following consequences, each of which illustrates different ways in which the role or task-based version of accountability loses its distinctive interpersonal address and expressive function:

- The moral reactive responses we have to each other (whether they be reactive attitudes, a lessening of trust, a breaking of relationship) are irrelevant to accountability, or icing on the cake at best.<sup>8</sup>
- The communicative/expressive aspect of moral accountability seems to be lost on this view of accountability, as there is no specific address to the person. There is merely a distribution of burdens or sanctions.
- The relations picked out by substantive responsibility are roles, not persons.
- Even if we aren't dealing strictly with roles but instead with action-types, trying to rule-ify accountability doesn't work in the end. Ultimately we are responding to people when it comes to accountability, not to mere thoughts, attitudes, choices, or actions.

Each of these consequences stem from the fact that the individual, with all their peculiarities, is not considered nor responded to. Accountability ceases to be something that happens between individual persons and becomes something that happens merely at the level of rules applied to role responsibilities. Substantive responsibility is grounded in interpersonal agreement, but its actual practice lacks the interpersonal character of ordinary moral accountability.

I propose that moral accountability should be understood as something that happens between persons and that has an expressive function. Accountability requires a specific address to the person being held accountable about the moral status of what they have done. This can be contrasted to substantive responsibility, which is something that happens between different roles or is determined by universal rules. It does not require any expressive function that pertains to the moral status of what the agent has done.

---

<sup>8</sup> Wolf (2015a) makes a similar point when critiquing Watson's notion of accountability (137).

Sanctions are still relevant for the interpersonal notion of moral accountability that I propose, but they must be sanctions that are tailored to the individual, taking into consideration who they are, what they have done, and how they are reacting to what they have done, among other contextual features. Even if the sanctions are given without moral blame, those sanctions must express some moral address to the person and be given from the participant stance.<sup>9</sup>

For example, say that my roommate has, yet again, failed to help me clean up the apartment, despite our mutual agreement for dividing up chores. I know that for her, chores represented punishment as a child and that she will likely feel outsized guilt if I remind her. Yet, the chores need to be done and the burden of housework needs to be shared. My response to her failure, however, should ordinarily be tailored to her. I might recognize her childhood history and remind her that I am not trying to inflict punishment upon her but to keep the house clean for the both of us. I might express my growing resentment but assure her that I don't really want her to feel guilty—I want her to contribute to the household. The point here is that whatever my response to her, in holding her to this expectation, I need to be responsive to my own feelings as well as to who she is and why she has flouted the expectation.

Holding accountable in this instance is an interaction between my roommate and me and implicates our particular relational dynamic.<sup>10</sup> It is not simply an abstract allocation of burdens on the basis of role responsibilities or universal principles. I might ordinarily feel indignation at her tendency to keep up with her duties for a week or so after being reminded but then to go back to failing to keep up her end of the bargain. Or, I might be frustrated at her refusal to problem solve and find better ways to help her share in the burden of cleaning. Whatever the details of the case,

---

<sup>9</sup> See section 2.4 in this chapter for a more in-depth discussion of Strawson's (1962) distinction between the participant stance and the objective stance.

<sup>10</sup> Depending on how healthy the relationship is, a good instance of holding accountable might attempt to alter the relational dynamic or reinforce it.

character is implicated in a way that goes beyond the basic notion of substantive responsibility, and holding accountable has a core expressive function that identifies the patterns of behavior and relating to each other that we are each engaged in.

Take another example in which you are confronting a cousin who once confessed to having feelings for you. Not only did the cousin have these feelings, but the cousin also chose to privately communicate those feelings while in a monogamous long-term relationship with someone outside the family. The vices exhibited by the cousin in this situation go beyond concerns about incest and include deceit of their partner, a failure to keep their own desires in check, and a violation of your trust, among other things. In expressing blame, one might mention these other things to fully characterize the extent of the moral transgression. What one chooses to express and how will be context dependent, but the response, whatever it is, will be to the unique circumstances of the moral transgressions that the cousin has committed, not to a violation of abstract moral rules. This response could include an attitude of blame, a breaking of the relationship with the cousin, or perhaps even an attitude like pity, if the cousin's formative circumstances were not conducive to developing a good character.

In both of these cases, accountability requires that one respond to the individual person, whether that be in one's blaming response that identifies the particular faults of the individual or in the ways in which one interacts with the other person to indicate that an expectation has been violated. Neither case reduces to role responsibility questions, though those are relevant for understanding both. The more general principles that govern roommate responsibilities and appropriate interactions among cousins help us understand the situations we find ourselves in, but we must look at the details of a particular case to gain a complete understanding of what has occurred. Likewise, we have general guidelines for how to hold others accountable in different kinds

of circumstances, but our responses will often need to be tailored to new and unique cases. We must exercise our practical judgment to do to the best that we can within our circumstances.

Substantive responsibility captures the general moral principles that should govern our collective conduct and provides a basis for institutions to hold their members accountable based on the responsibilities that attach to their various roles. But it does not capture key features of our interpersonal practices of holding each other accountable.

#### 1.4 Moral Accountability Requires Moral Attributability

As we have seen, accountability requires that one respond to the individual person and that any sanctions or responses express to that person morally significant features about what they have done. Hence, accountability necessarily invokes questions of character or practical identity. One's practical identity is implicated both at the level of what one has done and at the level of holding responsible. The character of those holding responsible and those being held responsible has implications for which moral reactive responses we have, which (if any) burdens or sanctions are applied, and how the moral address is communicated. So, moral accountability requires moral attributability.

This conclusion, that accountability requires attributability, is also supported by considerations of fairness—whatever blaming reactions, sanctions, interpersonal relational changes, or other moral responses we have to others should only express that wrongdoing has been committed if wrongdoing has in fact been done. Likewise, to make fair judgments between different agents, our judgments should reflect the individual circumstances that inform what each agent has done. If we merely assign burdens according to role responsibilities or action-types, we will miss peculiarities that require subtly different responses.

## 1.5 Delimiting the Question

In what follows, I will be concerned fundamentally with *character* understood within our practices of moral accountability in order to give an account of moral attributability understood within the context of our lives, our relationships, and the situations in which we find ourselves. While this is a project that privileges attributability, it is by no means disconnected from our practices of accountability. The assumption is that responsibility for character goes beyond judgments that conduct is “shoddy” or sub-par and that we are liable to blame, praise, other reactive attitudes and even some sanctions and burdens for our character.

In practice, there is not always a sharp line between attributability and accountability. We often find ourselves reacting first and discovering why later, and our judgments and reactions sometimes have very little space between them. Reactive blame and judgments of character may have but a hair’s breadth between them. However, the distinction between attributability and accountability is still relevant for this project, because my aim is to provide a framework for thinking about character itself, not to detail an additional landscape of reasons that govern the appropriateness of blame, such as standing. The core project is to give a compelling account of how to think about character that illuminates our ordinary practices of moral responsibility.

## Part 2: Responsibility in Context

In this section, I demonstrate that views of responsibility that focus on the fairness of blame to the person being blamed almost necessarily make responsibility out to be merely a burden. In an attempt to reclaim the idea of responsibility as something that is not merely burdensome, I situate moral responsibility within a relational context and expand the reactive emotional repertoire for accountability beyond blame.

## 2.1 A Range of Starting Questions

What should be our starting point for thinking about attributability and accountability in the moral domain? The answers given in the literature trend towards:

- cases where blame and/or sanctions are appropriate and/or deserved,<sup>11</sup> and
- when it makes sense to demand an explanation or justification for an attitude or behavior.<sup>12</sup>

Each of these starting points invoke questions about the fairness/appropriateness of accountability.

Most contemporary philosophers start with some version of the following question: “What conditions, capacities, or kinds of psychological states would make it fair/appropriate to blame someone, impose sanctions, or otherwise demand either justification or restitution?”

Little, if any, attention is given to the question: “When is it fair/appropriate to expect others to express blame, impose sanctions, or otherwise hold the offender responsible?” The burden of responsibility is treated as if it only falls on the person being held responsible, even though it is also frequently an unpleasant and difficult task to hold others responsible.

Both questions assume that responsibility is a burden. But is this the right way to think about responsibility?

## 2.2 Common Pitfalls

One of the key problems with these starting questions is that the practice of responsibility goes beyond blame and the distribution of sanctions; holding responsible also includes praise, gratitude, love, and the distribution of benefits.<sup>13</sup> A second issue is that even if we expand our notion of responsibility beyond blame, we can easily remain stuck in a blame and praise calculus for

---

<sup>11</sup> See Wallace (1994) and, as I’ve discussed in the previous section, Watson (1996).

<sup>12</sup> This is the general answerability approach endorsed by Scanlon (1998) and Smith (2015a).

<sup>13</sup> I will set aside questions about the asymmetry thesis of praise and blame—the idea that blameworthiness requires avoidability but praiseworthiness does not (Wolf 1990, Watson 1996).



deserved social esteem. (This is not necessarily an issue for the starting questions above, but they can easily allow for this kind of mistake, especially when our primary language of praise and blame lends to the idea that they exist on a single spectrum of best to worst judgments and responses.) If we don't understand the point of praise and blame, they can easily become hollowed out positive and negative points that get added together to form some overall total score that represents character.

There are at least three layers of mistakes in the move to score total praiseworthiness or blameworthiness: 1) It's tempting but wrong to think that responsibility just concerns the evaluation/assessment. Responsibility isn't like assessing beauty or giving a test score; our moral evaluations have direct relational implications. 2) It's overly reductive to think that character traits are related to one another, such that they're additive and comparable, rather than a complex whole that can't be reduced to its parts.<sup>14</sup> 3) Our language of the reactive attitudes is rich and varied, and multiple conflicting attitudes might reasonably apply in a given case—they don't cancel each other out. 4) This whole way of thinking wrongly assumes that there's a relationally independent fact of the matter, and it's not obvious that that's the case.<sup>15</sup>

This last point is related to a third problem: assuming that responsibility must be understood from an objective, god's-eye view that takes a full view of the agent in question. These wrong-headed questions about responsibility might go something like "In a final judgment in which the two options are eternal happiness or eternal damnation, does this person's entire life merit the former or the latter?"<sup>16</sup> However, our actual questions about responsibility and character make more sense

---

<sup>14</sup> As Michelle Mason (2003) notes, the object of contempt is generally the person as a whole and not isolated acts.

<sup>15</sup> Even if we are responding to the acts of historical figures or fictional characters, our parasocial relationships (or one-sided relationships) to those figures still inform our responses to them. These relationships are not merely trivial in the sense that we are related in some way to everything; they can be deeply personal. Think about the kind of emotional attachment that audience members feel to their favorite internet celebrity vlogger or the ways in which our responses to historical figures are frequently guided by our complex contemporary relationships with real people.

<sup>16</sup> An example of this in the philosophical literature can be found in Galen Strawson's paper "The Impossibility of Moral Responsibility" (1994): "As I understand it, true moral responsibility is responsibility of such a kind that, if we have it, then it makes sense, at least, to suppose that it could be just to punish some of us with (eternal) torment in hell and reward others with (eternal) bliss in heaven" (9).

within the immediate, context-sensitive features of a relationship. They might be more like “I keep getting frustrated at her when she gets into new, toxic relationships and asks me to go along with everything, even though the relationships never end up well. Should I stick by her as a friend, confront her, or distance myself from her?” The ordinary and everyday varieties of responsibility matter just as much as (and perhaps even more than) cosmic views of responsibility.<sup>17</sup>

When we make these mistakes (which are all, in fact, mistakes I have made throughout the course of thinking about this topic), I think we’ve lost sight of why responsibility matters.

### 2.3 Why Does Responsibility Matter?

There has been a growing trend in recent years to think that responsibility matters because it makes certain kinds of relationships possible, often with a focus on the values of mutual recognition and respect.<sup>18</sup> I think that something like this is the right approach, so long as we expand our reactive emotional vocabulary beyond blame.

In *The Second-Person Standpoint* (2006), Stephen Darwall argues that in order to see ourselves as responsible and standing in an intersubjective relationship of mutual recognition and respect with others, we need to be able to recognize each other’s competence, authority, and freedom as fellow reasoners. For Darwall, “second-personal reasons of the kind we are interested in ... are relational all the way down. They ultimately derive from normative relations that reciprocally recognizing persons assume to exist between them” (60). Responsibility and accountability are inextricably tied to recognizing each other as moral agents with the ability to make claims on each other.

---

<sup>17</sup> Interestingly enough, there’s been a trend in more recent theologizing that makes out responsibility relations with God to be more like responsibility relationships in a personal, human relationship.

<sup>18</sup> Strawson’s (1962) relational view of responsibility has informed much of the contemporary philosophical literature, including Scanlon (1998), Wallace (2019), Darwall (2006), and Fischer and Ravizza (1998).

Similarly, in *The Moral Nexus* (2019), R. Jay Wallace emphasizes the value of interpersonal recognition, which “involves two components: acknowledgment of the standing of other persons, as individuals to whom consideration is owed; and the ability to justify oneself specifically to those individuals, insofar as they might be affected by the things we do” (163). This roughly amounts to a moral regard condition and an answerability condition on responsibility.

Both Darwall and Wallace give compelling, relational pictures of moral responsibility that capture the importance of respect and regard and that point strongly toward expansive, universal moral rights. If responsibility is a burden that we must pay to access these goods, then that burden might be worth it.

At the same time, both Darwall’s and Wallace’s views leave me cold at some level.<sup>19</sup> Our relationships are much more diverse and wide-ranging than the basic relationship of mutual respect, and the contexts in which we tend to feel resentment and gratitude the strongest are in our friendships, romantic relationships, and families. Scanlon’s (2008) view of blame brings us a bit closer to these more personal contexts: “to blame a person for an action, in my view, is to take that action to indicate something about the person that impairs one’s relationship with him or her, and to understand that relationship in a way that reflects this impairment” (122-3). A praiseworthy character is also an integral part of why we become friends with someone, fall in love, or choose to remain in contact with our families. Abramson & Leite (2011) contend that reactive love paradigmatically involves responding to “particular kinds of morally laudable features of character expressed by the loved one in interaction with the lover” with acts of goodwill (673). If responsibility makes these kinds of relationships possible as well, then I can start to see why it’s really worth it.

---

<sup>19</sup> Darwall has since published a series of articles on [second-personal attitudes of the heart](#) to expand second-personal attitudes beyond mutual respect and recognition and address this kind of critique. For another critique of Kantian views that prioritize respect instead of love, see Brewer (2021).

Responsibility is indeed sometimes a burden, but if we understand it within its rich relational context, we also see that responsibility can be a joy. The prize we get for being responsible and developing a good character is being able to have meaningful and rewarding interpersonal relationships, not merely being able to avoid sanctions.

#### 2.4 The Core of Moral Responsibility

If moral responsibility is relational and closely tied to the project of living well with ourselves and others, then what is its scope? What separates moral responsibility from other kinds of responsibility?

Like most other contemporary authors, I will be taking a Strawsonian approach to this question. My approach will be informed primarily by P. F. Strawson's (1962) following remarks:

I want to speak, at least at first, of something else: of the non-detached attitudes and reactions of people directly involved in transactions with each other; of the attitudes and reactions of offended parties and beneficiaries; of such things as gratitude, resentment, forgiveness, love, and hurt feelings.... I have mentioned, specifically, resentment and gratitude; and they are a usefully opposed pair. But, of course, there is a whole continuum of reactive attitude and feeling stretching on both sides of these and—the most comfortable area—in between them.

Like Strawson, I hold that gratitude and resentment are paradigmatic instances of holding responsible. I will not prioritize resentment over gratitude or gratitude over resentment.<sup>20</sup> I also take it that there is a wide range of reactive attitudes between and around the paradigmatic pair that admit to moral responsibility. The valences of these attitudes are complicated and go beyond how negative or positive they are. While gratitude and resentment are core to moral responsibility, they need to be fitting or appropriate. You can wrongly feel resentful or grateful towards someone. I do not think,

---

<sup>20</sup> Michael McKenna (2012) similarly works to expand the concept of moral responsibility beyond blame for failing to comply with obligations, but McKenna takes the case of appropriate blame to be a central, paradigmatic case that others can be extrapolated from (195-204). I will employ a similar strategy but place both appropriate resentment and gratitude as central cases of moral responsibility.

however, that there is a single reductive rationale that makes the reactive attitudes appropriate. This picture does not allow me to draw a clean line around the bounds of moral responsibility, but it does allow me to gesture at core instances and make an argument for why other instances may reasonably fall under the scope of moral responsibility. It also makes room for the rich emotional vocabulary we have within the participant stance.

Again, drawing from Strawson, I take it that the participant stance broadly demarcates the stance by which we hold others responsible and reason with them as equals who can do the same:

If your attitude towards someone is wholly objective, then though you may fight him, you cannot quarrel with him, and though you may talk to him, even negotiate with him, you cannot reason with him. You can at most pretend to quarrel, or to reason, with him.

What this suggests is that the participant attitudes are only appropriate in relationships in which the two parties can reason with each other and mutually contest and shape the norms and expectations that bind them. At the same time, we sometimes choose to take up the objective stance to deal with the strains of interpersonal involvements. Deploying the participant stance and the mode of character is an active choice in some cases.

There is a question on this view: are we responsible for anything that is morally neutral that does not elicit emotions in the obvious range around resentment or gratitude? Fischer and Ravizza (1998) claim that their “Strawsonian view of moral responsibility allows for moral responsibility for ‘morally neutral’ behavior. For instance, one can be morally responsible for simply raising one’s hand (where this is not a signal or in any way morally significant)” (8, Footnote 11). This is because “someone is a morally responsible *agent* insofar as he is an appropriate candidate for at least some of the reactive attitudes on the basis of at least some of his behavior (or perhaps his character)” (6). This does not require the actual application of a reactive attitude (8).

Fischer and Ravizza’s response seems to be that liability to responsibility generally allows for responsibility for morally neutral behavior with no clear moral significance. I suspect, however,

that reactive feeling is more central to the participant stance than Fischer and Ravizza's view licenses, and if gratitude and resentment are the paradigmatic pair, then it seems that morally neutral reactions might be less central to moral responsibility.

Take the final remark in the quote from Strawson given above: "there is a whole continuum of reactive attitude and feeling stretching on both sides of these and—the most comfortable area—in between them." What is the reactive attitude and feeling that occurs when I have a normally polite but tired interaction with a cashier at the gas station down the road? Is it the most comfortable area in between gratitude and resentment? Is minimal respect and fellow feeling on the continuum of reactive attitudes? Plausibly so. Is it a clear-cut case of responsibility? Probably not.

As these kinds of interactions have little effect on our relationships with others, I will leave it an open question as to whether they should be included as cases of moral responsibility. In what follows, I will focus on responsibility reactions of gratitude and resentment and other more closely related reactive emotions to narratives of character within interpersonal contexts.

## Conclusion

In the next chapters, I build my theory of character, which helps us identify who we are as actual, concrete individuals within interpersonal moral relationships. Before I delve into character as a whole, I turn in the next two chapters to responsibility for choice and judgment, respectively, to explain how my view relates to other contemporary views of responsibility and to motivate why I want to include both judgment and choice as elements of narratives of character.

## Chapter 2: Wallace's Choice-Centered Account of Moral Responsibility

How central is control to responsibility? Are we responsible for anything outside of our direct control?

R. Jay Wallace's (1994) choice-centered account of moral responsibility is notable for its emphasis on direct control as the defining condition of responsibility. As I argue in Chapter 3, I think that a choice-centered view of this kind gets something right, namely that choice is morally relevant for responsibility and cannot be reduced to Scanlon's (1998) notion of judgment-sensitive attitudes, or "those attitudes that, in a rational creature, should be 'under the control of reason'" (272). In this chapter, I argue that choice alone is insufficient to capture our judgments about what we can be held morally responsible for. While choice holds *a* central place in our understanding of moral responsibility, it does not hold *the* central place.

First, I present Wallace's account. I then explain where I depart from Wallace and why. Finally, I address concerns about moralism. Through this commentary I will begin to sketch the contours of my positive view, which I will present in Chapter 5.

### Part 1: Wallace

Wallace presents two distinct arguments that place choice and control at the center of what it is to be responsible. The first argument isolates the qualities of will that individuals can be responsible for; the second identifies the general capacities necessary for responsibility.

#### 1.1 Wallace's Larger Project

Before I discuss Wallace's arguments for his choice-centered view of moral responsibility, it is worth painting a general picture of Wallace's overarching project. Wallace's project is primarily to give an account of moral responsibility that begins with our moral practices of holding accountable.

Wallace's starting point "is not the question, What is it to be a morally responsible agent, but rather, What is it to hold a person morally responsible?" (1).

Wallace begins by analyzing the basic stance of holding someone to an expectation, which he defines in terms of the negative reactive attitudes, such as resentment, indignation, and guilt (21).

For Wallace, "to hold someone to an expectation is essentially to be susceptible to a certain range of emotions in the case that the expectation is not fulfilled, or to believe that the violation of the expectation would make it appropriate for one to be subject to those emotions" (21). These "emotions" are, of course, the reactive attitudes as Wallace defines them.<sup>21</sup>

From this account of holding someone to an expectation, Wallace provides an answer for what it is to hold someone morally responsible: "To hold someone morally responsible is to hold the person to moral expectations that one accepts" (63). The main question that motivates Wallace's picture of what it is to be a morally responsible agent is What conditions make it fair to hold someone morally responsible? (1).

## 1.2 Wallace's First Argument

Wallace's first argument for what it is to be responsible moves from our practice of placing moral demands and expectations on each other to the qualities of will that can be regulated by those demands and, subsequently, the qualities of will that we can be responsible for. This argument doesn't yet answer questions about the fairness of holding responsible; it simply identifies the qualities of will targeted by our moral practices of holding responsible. This will be necessary for answering the broader question about fairness in the second argument. The first argument proceeds as follows (this is my own reconstruction):

---

<sup>21</sup> Wallace takes a narrower view of the class of reactive attitudes than Strawson (1962), focusing in on a range of negative attitudes (33-40). Strawson, by contrast, recognizes a "whole continuum of reactive attitude and feeling stretching on both sides of" and between "resentment and gratitude."



1.P1. “[M]oral obligations are supported by reasons, of the sort that can be expressed in the form of principles” (129).

1.P2. Our practice of holding people to moral obligations requires a) that there is some possible justification/set of reasons that one could give in support of those obligations and b) that the justification/set of reasons “could move others who accept those reasons to comply with such obligations themselves” (129-130).

1.P3. “The reasons expressed in moral principles are *practical* reasons, or reasons for *action*” (132).

1.P4. “This suggests that such reasons should regulate what we do, the bodily movements that we make... insofar as they manifest a choice that is made by the agent” (132). This is because only our choices can be directly influenced by the reasons in moral principles (132).

1.C. So, the qualities of will that are presupposed in our practice of making moral demands, and subsequently that we are responsible for, are our choices.<sup>22</sup>

Roughly, this argument begins with the idea that the reasons expressed by our moral obligations are reasons for action. Holding people responsible for the violation of those obligations requires that we can jointly communicate about the reasons underlying those principles and be motivated by those reasons. Wallace concludes that only our choices can properly mediate between the reasons expressed in moral obligations and our actions, allowing us to be motivated by the reasons that support those obligations. Hence, the qualities of will picked out by our practices of moral responsibility are our choices.

More clarification is needed, however, to fully understand why Wallace’s argument picks out choice and not other phenomena such as emotions or feelings. Wallace supplements premise 4 by distinguishing between phenomena that can be influenced directly by reasons and phenomena that can be influenced only indirectly by reasons (131-132). The picture of our practices of holding others to moral obligations presented in the above argument “suggests that such obligations can properly be focused only on phenomena that are susceptible to being influenced directly by reasons”

---

<sup>22</sup> The language of “quality of will” is taken from Strawson, who, unlike Wallace, speaks of it primarily in terms of regard towards others, including good and ill will, as well as indifference or lack of concern.

(131).<sup>23</sup> For Wallace, only choices can get us the direct influence that is required for moral responsibility. Emotions, feelings, and mere bodily movements, by contrast, “are not the sorts of states that can directly be controlled by the reasons expressed in moral principles: such states as love, esteem, and goodwill are generally not states that could be produced simply by the belief that there are moral considerations that make them obligatory” (132). States of emotion and feeling might still be indirectly influenced by moral justifications, but this occurs through the mediation of our choices—if one believes, for instance, “that goodwill is morally desirable, and pride or patriotism morally objectionable, one might take steps to cultivate the one emotion and to rid oneself of the others” (132, Ftnt. 21).

In sum, the first argument begins with our practice of giving moral demands to one another, which must be justified by reasons that we can explain to each other and that can motivate us to comply with those demands. These reasons are reasons for action, and so the only things that can be directly influenced by these reasons are our choices. At best, our emotions can only be indirectly influenced by those reasons through the mediation of our choices. Hence, our moral demands target our choices, and those are the qualities of will that we are responsible for.<sup>24</sup>

### 1.3 Wallace’s Second Argument

Wallace’s second argument directly addresses the main question of his project: When is it fair to hold people morally responsible? If the first argument holds, then, according to Wallace, “it will be reasonable to adopt this stance [the stance of holding people to moral obligations] only towards people who possess what I refer to as powers of reflective self-control: the general ability to grasp

---

<sup>23</sup> As the second argument brings out, Wallace thinks that indirect influence isn’t sufficient for the control needed to be able to avoid sanctions and meet basic requirements of fairness.

<sup>24</sup> One might wonder, if responsibility comes down to choice, how does Wallace deal with responsibility for negligence and recklessness? On Wallace’s view, negligence may be traced to an earlier fault in choice and recklessness involves making a choice while being aware of the risk associated with that course of action (138-139).

and apply moral reasons and to regulate their behavior by the light of such reasons” (154-155). I have reconstructed the second argument as follows:

2.P1. The conditions of responsibility “must make it fair to demand that *s* comply with moral obligations we accept, where the obligations are supported by distinctively moral reasons” (157).

2.P2. It is unfair to hold someone to moral obligations one accepts if the agent “lacks the basic power to do the sort of thing that we demand of him” (161). This is because they will “lack the basic capacity to avoid the sanctions to which violating a demand normally exposes one” (162).

2.P3. If someone lacks either the capacity to grasp and apply moral reasons or the capacity to regulate their behavior by the light of such reasons, then they lack the basic power to comply with moral obligations we accept.<sup>25</sup>

2.C. So, it is only fair to hold someone accountable if the person has “powers of reflective self-control: the general ability to grasp and apply moral reasons and to regulate their behavior by the light of such reasons” (155).

This argument proceeds from the idea that it’s unfair to hold people to moral obligations and expose them to the possibility of sanctions if they can’t avoid those sanctions. Wallace concludes that to be able to comply with our moral obligations and avoid sanctions, people must have the ability to act in accordance with our moral expectations. This, in turn, requires that they be able to recognize the reasons expressed in moral principles and control their behavior to conform to those reasons.

For Wallace, the general ability to grasp and apply moral reasons requires that one must be able to “bring the principle to bear in the full variety of situations to which it applies, anticipating the demands it makes of us in those situations, and knowing when its demands might require adjustment in light of the claims of other moral principles” (157). This ability requires an understanding of the moral concepts and values involved in justifying moral principles, as well as background abilities of attention, concentration, and judgment in order to be able to deliberate and

---

<sup>25</sup> Wallace claims that in the absence of either of the two powers, “people will be fundamentally unable to understand why they should comply with moral obligations; or if they can understand this, they will be fundamentally unable to translate their understanding into action and thereby to avoid the responses of moral sanction” (162).

apply those concepts to potential courses of action (157-158). It does not require that one accept all the moral obligations that we do (163). People may, and often do, have the “*ability* to grasp the reasons that support our moral obligations, and to regulate their behavior in light of such reasons, even if they have rejected those obligations in fact” (164).

In order to have the ability to regulate or control one’s behavior in light of reasons, one must have “a capacity for critical reflection: the ability to step back from one’s immediate desires and assess the actions they incline one to perform, in light of the moral reasons one has grasped and accepted” as well as “the capacity to make choices as a result of deliberation” (158).<sup>26</sup> The ability to grasp and apply moral reasons is thus necessary to have the ability to regulate or control one’s behavior in light of reasons.

These capacities—the general ability to grasp and apply moral reasons and the ability to regulate one’s behavior by the light of such reasons—are not necessarily all or nothing, nor are they sharply divided. They admit to various levels of development which may allow for at least partial accountability, though the possession of these basic powers is required for responsibility (167).

Additionally, either of these two powers could be limited in different ways:

One’s ability to appreciate and focus on the moral reasons in favor of an obligatory course of action may be reduced (either in general, or for a limited period of time); or internal conditions of emotion and desire may render it difficult to translate one’s moral choices effectively into action (160).

If either of these powers are sufficiently limited, then the agent cannot be fairly held responsible.

In sum, on Wallace’s view, in order to fairly hold someone responsible, it must be reasonable for us to demand that that person comply with moral obligations we accept (161). If the agent

---

<sup>26</sup> Wallace thinks that “it is necessary that the agent should have some capacity to translate her choices into behavior” (158). This does not imply that one must always be able to translate one’s choices into action in order to be responsible. Take the following example: One might choose not to take a friend to the airport (having previously promised to do so) and then discover that one is physically incapable of following through on one’s promise. Yet, “[i]n these cases, despite the presence of physical constraints, one’s omission nevertheless expresses precisely the kind of choice that our moral obligations prohibit” (142).

cannot do the sort of thing that we demand of them, then they lack the ability to avoid moral sanctions. And, “it is unreasonable to hold the person to moral obligations under these conditions” and potentially expose them to the harms of moral sanction (161). Holding such a person responsible would thus be unfair. Both the capacity to grasp and apply moral reasons and the capacity to regulate one’s behavior in light of one’s reasons are necessary to have the basic capacity to avoid moral sanctions. So, it is only fair to hold someone accountable if they have the two “powers of reflective self-control” developed to a sufficient degree.

#### 1.4 Summary

To summarize both arguments, on Wallace’s view, the states we can be held responsible for are our choices, and we can only be held responsible for these states if we possess the ability to understand and apply moral reasons as well as the ability to control our behavior in light of those reasons. So, to be a morally responsible agent, one must be able to reason within the moral domain and regulate one’s behavior in light of one’s reasons; however, what we care about when we are holding others responsible is their choices, not their reasons. Their reasons are relevant primarily for providing the background conditions necessary in order for them to be responsible for their choices that are expressed in action. As such, Wallace’s account of being responsible is radically choice-centered.

Before I begin my discussion of Wallace’s view, there is a very real worry that Wallace and I are talking past each other. Wallace makes several remarks that suggest he strongly separates questions about complying with moral obligations and questions about character. For example, Wallace notes that moral obligations tend to track actions rather than reasons:

Note also that when we hold other people to moral obligations we accept, we do not literally demand that they comply with the obligations because they grasp the moral reasons for doing so. At least for purposes of apportioning blame, we generally do not care so much

why people comply with the moral obligations to which we hold them, so long as they do comply with those obligations in fact (130-131).

In a footnote attached to the previous quote, Wallace suggests that we do care about reasons when thinking about character:

It is of course otherwise when questions of character assessment are at issue: there we do care very much what an agent's reasons for acting really were. There may also be a disanalogy between the first- and third-person cases, with respect to questions of blame. Thus, though it would be strange for others to resent or be indignant with me when I comply with moral obligations for the wrong reasons, it would not necessarily be strange for me to feel guilt in such cases (Ftnt 19).

It appears, from these quotes, that Wallace and I are simply concerned with different things. Wallace is concerned with moral obligations, and I am concerned with character.

However, I think we have a more substantive disagreement. I do not find it strange to resent or be indignant with others who comply with moral obligations for the wrong reasons, especially if those reasons are significantly wrongheaded and not garden varieties of self-centeredness. I imagine that Wallace is in the background considering a distinction something like Watson's (1996) distinction between attributability, which has to do with character evaluation but does not yet warrant sanctions, and accountability, which raises questions of avoidability and fairness due to its connection to sanctions. Wallace seems much more concerned with the latter rather than the former. As I have argued in Chapter 1, I think that moral accountability requires moral attributability. Character should not be relegated to a toothless kind of moral responsibility that loses the full range of the interpersonal reactive attitudes and responses, both positive and negative.

There is some question as to where Wallace's notion of moral responsibility falls between accountability and substantive responsibility on my taxonomy, as his focus on moral principles and obligations suggests something more like a rule-based or role-based conception of responsibility. However, the reactive attitudes at the core of his view have a clear expressive function, even if they are only targeted at an agent's choices and not their reasons. Assume for now that we are both

talking about accountability, due to our common commitment to the centrality of the reactive attitudes. I think the best way to categorize my conflict with Wallace is as a disagreement about what we can be held accountable for. It is my view that we can be held accountable for our character, not merely our choices, and that our practices of accountability should allow reactive responses to an agent's patterns of reasoning. Likewise, it seems to me that our moral principles do pick out certain features that implicate character, including what reasons we should recognize and under what circumstances we should perform certain actions. Gratitude can quickly turn to resentment when we discover that someone's good deeds done towards us were only done for opportunistic gain.

As I will show in the next section, Wallace's focus on choice as the central object of responsibility cannot capture the breadth of our ordinary practices of holding others responsible.

## Part 2: Departures from Wallace

In what follows, I present some critiques of Wallace. While I think he too narrowly focuses on the negative reactive attitudes and on choice, I think he's right to gesture at an important distinction between direct and indirect control. It turns out, however, that *both* direct and indirect control matter for responsibility, just in different ways.

### 2.1 Departures from Wallace's First Argument

Let's start with some departure points from the first of the two arguments I reconstructed in Part 1.

Take Wallace's first argument. The first premise worth discussing is 1.P3. "The reasons expressed in moral principles are *practical* reasons, or reasons for *action*" (132). Depending on how

the word ‘action’ is interpreted, it is unclear to me that this is true for all moral principles.<sup>27</sup> On Wallace’s view, these reasons for action regulate our “bodily movements insofar as they manifest a choice that is made by the agent” (132).<sup>28</sup> If action requires choice, then 1.P3 does not capture the breadth of reasons given by moral principles—some of these reasons seem to govern our judgments, our attitudes, and other states that are not the result of our choices.

Take Angela Smith’s (2005) example of forgetting a good friend’s birthday:

I forgot a close friend’s birthday last year. A few days after the fact, I realized that this important date had come and gone without my so much as sending a card or giving her a call. I was mortified. . . . But what, exactly was the nature of my fault in this case? After all, I did not consciously *choose* to forget this special day or deliberately *decide* to ignore it. I did not *intend* to hurt my friend’s feelings or even *foresee* that my conduct would have this effect. I just forgot. It didn’t occur to me. I failed to notice. And yet, despite the apparent involuntariness of this failure, there was no doubt in either of our minds that I was indeed, responsible for it (236).<sup>29</sup>

Smith seems right both that failing to remember could not be traced back to some choice and that she was responsible for forgetting her friend’s birthday. She had violated a standing expectation in that relationship, and her failure to remember suggests that her friend’s place in her general evaluative outlook had somewhat diminished.<sup>30</sup> Even if no choice was involved, it still seems that Smith was responsible for the underlying attitudes that explained her forgetting. Furthermore, it

---

<sup>27</sup> Nothing about this disagreement hangs on a particular understanding of “moral principles” as opposed to “moral ideals” or “moral standards.” Wallace defines moral principles as “propositions that isolate the features of situations or actions that justify particular claims of moral right; they specify the reasons that a given course of action is obligatory or prohibited” (109). While this narrows the class of what our moral principles govern (on which I do disagree with Wallace), the language of principles picks out a broad category that might be equally well referred to as “expectations,” “standards,” “ideals,” etc.

<sup>28</sup> This also includes inaction. As Wallace illustrates: “Take the obligation of nonmaleficence, not to harm people in the ordinary pursuit of one’s ends. In accordance with the remarks just made, this must be construed as an obligation not to make bodily movements that harm someone, as the result of a choice to bring about such harm” (133).

<sup>29</sup> Assume that, in this case, there was a standing expectation that each friend remember the other’s birthday and mark it in some way, as an expression of a tightknit friendship.

<sup>30</sup> If you thought that it seems like a jump from merely forgetting something once to the idea that the friend has slipped in Smith’s evaluative outlook, you’d be right. Assume that it’s plausible enough for now that at least some non-voluntary behavioral patterns can serve as the basis for reasonable interpretations of evaluational frameworks. This is a more complicated task than Smith tends to recognize, as I’ll argue in Chapter 3.



would be appropriate for the friend to respond with negative reactive attitudes (though, in this one-off case, disappointment may be more appropriate than resentment).

We also care deeply about the reasons why others conform to our expectations. Even if someone acts rightly, we might justifiably be upset if it turns out that their actions are motivated by selfish concerns rather than by genuine care for others. For example, if someone has been acting as lovingly as possible in a relationship, but it turns out that they are only doing so for their partner's money, their partner may be rightly hurt and angry. Or, if I find out that someone has gifted me a hat that they hate to get rid of, when they told me it suited me perfectly, I would be rightly miffed.

Finally, it seems plausible that moral principles might govern our emotions and feelings themselves, absent any (or even contrary to) choices made by the agent. Think about, for instance, the moral critiques of *schadenfreude* (feeling pleasure at the pain of others) or the virtue theoretic insight that right action also requires right feeling. It seems problematic, for instance, if someone feels disgust at fat bodies even if they explicitly avow that fat bodies are beautiful. Perhaps she never chose to inculcate these attitudes in herself, but her environment was strongly fatphobic. Assume that, despite the agent's best efforts, she slips up when she's tired and makes a visibly disgusted face at a fat woman getting out of the pool. That may happen at a non-voluntary, reactive level in the absence of any choice.<sup>31</sup> If the woman is her friend, the reassurance that she disavows those feelings

---

<sup>31</sup> Hieronymi (2019) delineates the non-voluntary from the involuntary: "So, belief is not active in the way an ordinary intentional action is, it is not voluntary, but it is not passive in the way that perception and sensation are—nor is it involuntary in the way digestion or blinking or accidents are. It is up to you in the way your answer to a question is up to you. Thus I will say that believing is non-voluntary. It is, I think, a non-voluntary activity. Many other states of mind—indeed, many of those we find most important—are also non-voluntary. I would include, in this wide category, intentions, decisions, and attitudes such as resentment, gratitude, trust, admiration, contempt, and satisfaction at a job well done. These are, like belief, attitudes for which we can be asked our reasons, but which we cannot adopt at will" (8). In what follows, I treat Robert Adams' language of the involuntary as interchangeable with Hieronymi's language of the non-voluntary.

As Hieronymi also helpfully notes, the non-voluntary is still up to us in an important sense: "Thus the fact that an attitude is non-voluntary, in the sense I have outlined, should not lead us to think it is not up to us or that we are passive with respect to it. These attitudes are up to us in the way our answers to questions are up to us—they express our take on things. Unlike your bone structure or birthplace, they change with your changing evaluation of what is true, good, worthwhile, or important. Unlike your headache or a visual illusion, it makes sense to ask you for your reasons for

and is doing her best to get rid of them may mean very little in the way of an excuse. These kinds of feelings and attitudes, even in the absence of choice, seem to matter a great deal to us in our interactions with others, and it is very common for our moral critiques of others to pick out the morally questionable attitudes that others express.

The second premise I reject is part of 1.P4: only our choices can be directly influenced by the reasons expressed in moral principles (132). This is not always the case. Sometimes reasons strike us and we come to see their value/truth without any conscious choice to accept that reason. In the same section, Wallace makes the further puzzling claim that “it is only through the mediation of our choices that the reasons expressed in moral principles may influence either our emotions or feelings, or the bodily movements we make” (132). But it often seems that our emotions, feelings, and even bodily movements can be directly influenced by moral reasons.

For example, I might come to see my own value and worth in response to someone else’s kind act towards me. I might recognize that there is nothing wrong with a gay acquaintance of mine, despite what I’ve consciously accepted from my religious upbringing. Or I might witness factory farming and be revulsed at the sight of the animals’ living conditions. Each of these examples seems to pick out a reason expressed in moral principles.<sup>32</sup> None of these examples involve a direct mediation of my choices from moral principles to my emotions or feelings.

In some sense, perhaps, there is an indirect mediation of my choices, as I would have made certain decisions that placed me in a position to recognize those reasons. But those choices were not necessarily choices to have those reasons influence my emotions and feelings. They might have simply been a choice to go out to the store, where someone interrupted my day unexpectedly with a

---

them. They are not simply out of your control, even though you cannot adopt them for any reason you take to show it worth doing” (11).

<sup>32</sup> Remember that Wallace defines moral principles as “propositions that isolate the features of situations or actions that justify particular claims of moral right; they specify the reasons that a given course of action is obligatory or prohibited” (109). Each of these cases involve seeing certain features of situations or actions that justify claims of moral right that have implications for whether certain courses of action are obligatory or prohibited.

kind act. Or a choice to attend university, where I encountered someone who was openly gay. Or a choice to open a social media app, where the video of factory farming showed up on my feed. There isn't the right connection between those choices and the emotions and feelings I develop for these emotions to be traced back to my choices.<sup>33</sup> It still seems that I am responsible for the attitudes themselves, independent from my choices.

The reasons in moral principles can also influence our bodily movements without the intervention of direct control. Take a case in which, without thinking, a mother throws herself in harm's way to protect her child. No deliberation or conscious choice is necessary to explain this action. All that is necessary is that the mother saw that her child was in danger. Harry Frankfurt (2004) motivates this idea nicely when he describes the "one thought too many" problem as it pertains to our expectations about love:

I do have problems with a couple of the details. For one thing, I cannot help wondering why the man should have even the one thought that it's his wife. Are we supposed to imagine that at first he didn't recognize her? Or are we supposed to imagine that at first he didn't remember that they were married, and had to remind himself of that? It seems to me that the strictly correct number of thoughts for this man is zero. Surely the normal thing is that he sees what's happening in the water, and he jumps in to save his wife. Without thinking at all. In the circumstances that the example describes, any thought whatever is one thought too many (36, Footnote 2).<sup>34</sup>

Not only is choice or deliberation not necessary to explain certain actions or bodily movements, it may also be odd for choice and even thought to enter into our reactions to certain situations at all. This does not mean, however, that the person's values, cares, and character are not still revealed through their bodily movements in these cases.

---

<sup>33</sup> I will say more about these kinds of cases when I discuss tracing accounts in Section 2.4 of this chapter.

<sup>34</sup> This example, as well as the earlier ones in this paragraph, represent instances of paradigmatically moral actions. If there is a worry that these examples are too partial or do not capture the impartial nature of morality, they could be readily adjusted to something like an off-duty lifeguard seeing a swimmer drowning and instantly diving into the water to help. In these cases, as in the cases above, it seems that we sometimes leap into action reflexively but that we are still responsible for these acts.

Each of these cases I have discussed are explained first by reasons-recognition, not explicit choice. Choice is important at a later stage, at which the person might endorse or reject that initial judgment or impulsive action. But it appears that in each of these cases, the reasons expressed in moral principles may directly influence our emotions, feelings, or bodily movements without the mediation of choice. Choice is not the only state susceptible to the direct influence of moral reasons.

## 2.2 Departures from Wallace's Second Argument

Move to the second argument. Here, the premise I reject is 2.P3. "If someone lacks either the capacity to grasp and apply moral reasons or the capacity to regulate their behavior by the light of such reasons, then they lack the basic power to comply with moral obligations we accept."

It's unclear to me that this is true for all moral obligations, assuming that the capacity to regulate behavior requires choice or direct control. This falls out of my rejection of 1.P3. If moral obligations and principles govern things outside of action, then we should revise 2.P3 to the following: If someone lacks either the capacity to grasp and apply moral reasons or the capacity to regulate their behavior by the light of such reasons, then they lack the basic power to comply with *all* moral obligations we accept. For a subset of our moral obligations, all that is required is the capacity to grasp and apply moral reasons.

Take an alcoholic who, when sober, just thinks about getting the next drink but, when drunk, desperately wants to be sober. Assume that, when drunk, the alcoholic wants to be sober because she recognizes how her actions are affecting herself and the people around her. Assume also that this is part of the early stages of the alcoholic realizing that she needs to deal with her alcoholism and that culminates in her seeking help after some time. While drunk and unable to exercise control, this person can grasp and apply at least some moral reasons and norms applying to her situation. The mere inability to regulate her behavior in this state does not remove responsibility

for her internal desire for sobriety. Even if she lacks direct control over her addiction, she still meets at least some of our expectations concerning the regard we should have for each other.

One might worry, if the ability to govern one's behavior is not necessary to be held responsible for violating at least some moral obligations, then how can an agent be able to avoid sanctions? Isn't that an unfair picture of moral responsibility? On my view, the agent can still avoid sanctions, but the avoidability does not have to come from the ability to choose otherwise—it can also come from the ability to respond to values in the absence of choice. If I should have seen the badness of some practice, like female genital mutilation (FGM) and I fail to see it despite being acquainted with the practice and having the ability to see that it was bad, I am still responsible. In that case, I could have recognized other reasons, even though I failed to do so.

What is also important in these cases, as a matter of fairness, is that our reactions of blame and sanctions should take into account how well someone reasoned or acted given what we could reasonably expect of them in a given context.

### 2.3 Conflicts Between Attitudes and Choices

Turn now to potential problem cases for choice-centered views like Wallace's, namely cases in which attitudes and choices are in tension with each other. These cases can be problematic for choice-centered views if they cannot account for the agent's reasons and how the agent's intuitive rational processes inform and often conflict with the agent's choices. Take the following case:

**A Tale of Two Activists:** Imagine that there are two people, both of whom are making the same choices to fight racial injustice. Agent A lacks racist attitudes, while Agent B still finds themselves with deep-seated racist attitudes that they repeatedly must struggle against to conform with morality. If these two agents are only morally responsible for their choices,

then there should be, in principle, no difference in our assessments of their choices. But it does seem like we would (and should) morally assess these people differently.<sup>35</sup>

In what follows, I will look first at how Wallace might handle this kind of case. I will then turn to another potential choice-centered solution.

What might Wallace say? One resource is suggested by Wallace's discussion of coercion. In that discussion, Wallace distinguishes between "choosing x" and "choosing x-rather-than-y." This is relevant to the case of coercion, because we might find "choosing x" to be impermissible but "choosing x-rather-than-y" to be permissible (144). For example, "stealing food" is generally prohibited by moral norms, but "stealing food to avoid starvation" is often permissible (145). It is important, then, to get the content of the agent's choice correct.

Perhaps the difference between the two agents, on a Wallace-type view, is as follows: The agent that lacks racist attitudes never has the racist actions enter in as possibilities into their motivational structure, so their choice is always "choose x," where x is some non- or anti-racist action. The agent that has racist attitudes always has to contend with some affective pull that those attitudes present, and so they are always "choosing x-rather-than-y," where x is some non- or anti-racist action and y is some racist action.

This is a promising expansion of Wallace's choice-centered view. It complicates the picture to account for not only the actional possibilities that are salient to the agent but also the

---

<sup>35</sup> One might think that the difference between these cases should be explained by the laudable force of will that the agent with racist attitudes exerts in order to act rightly. This is a tempting answer, though it's unclear to me that this explanation is a desirable one. It is true that the agent is making the right choices despite their attitudes, and that is something we should want to preserve in our explanation of the case. But there is a difference between praising someone for making the right choices and praising someone for exerting force of will.

We might have reason to worry about the popular tendency to praise discipline of choice, where one constantly resists certain desires (to eat sweet or fatty foods, for instance) by exerting a strong will to do something else (to lose weight). By fetishizing this kind of self-discipline, we lose sight of the greater aim of living a more balanced life in which one's desires and choices are more aligned. It would not do to fall into a moral authoritarianism that denies the self and the things that help one flourish. While sometimes self-discipline is necessary and good to achieve a more integrated life, excessive self-discipline can backfire and prevent us from addressing the root issues that are causing a mismatch between our desires and ends.

motivational strength of the reasons the agent has. This version of the view can better account for who we are as complex and often conflicted human beings. One difficulty for this view, however, is that it must answer questions about how the agent's reasons and the agent's motivations fit together. This is not an impossible task, but it is not particularly relevant at the level of ordinary interpersonal responsibility and accountability.

My own worries about this view fall into a different category. The main worry is that this view still tries to place everything that we are morally responsible for within the realm of what we can directly control. Even the revised view doesn't capture the conflict and interplay between the voluntary and non-voluntary parts of ourselves, which seems necessary for understanding who we are as human moral agents. This echoes Robert Adams' (1985) likening of the soul to the state:

[T]he order of the soul is not a pure democracy, and certainly not mob rule, but something more like the American system of representative government with 'divided powers,' with opposing tendencies and competing interests retaining an independent voice and influence .... It is important for the individual, as for the state, to be able to *act* fairly consistently over time in accordance with rationally coherent policies subjected to ethical reflection. But it is also important for the individual, as for the state, to have potential sources of dissent within .... The ever present possibility of internal conflict is not only a vexation and a potential hindrance to resolute action; it is also a wellspring of vitality and sensitivity, and a check against one-sidedness and fanaticism (10-11).

If we lose responsibility for these sources of dissent, we lose responsibility for instances in which our desires and feelings point us in a better direction than our considered evaluative outlooks.

Nomy Arpaly (2003) highlights the moral importance of cases of what she terms "inverse akrasia" or "doing the right thing against one's best judgment" (9).<sup>36</sup> These are cases in which internal sources of dissent guide the agent towards what is virtuous against the agent's endorsed views. Think about a character like Han Solo who ostensibly cares only about his own profits and the wellbeing of his own enterprises and yet consistently acts against his self-interest to help Luke and other members of the rebellion. This is plausibly explained by some level of recognition of what

---

<sup>36</sup> For an earlier discussion of inverse akrasia, see Arpaly and Schroeder (1999).

is good, even if, in Han's best judgment, he should be prioritizing his own interests. It seems that we want to preserve responsibility for these non-voluntary wellsprings of sensitivity, especially if they can guide us in the right direction and towards right action. This holds even if we in fact do the wrong thing in line with our reflectively endorsed views. In some cases, this can make the judgment worse: "you knew on some level!" "you could have done the right thing but you still chose wrongly!"

A secondary worry is that, even on this expanded view, Wallace cannot account for responsibility for the reactive attitudes themselves. This is a strange result, as Wallace's project starts with the reactive attitudes. It seems, too, that we tend to hold others responsible for their reactive attitudes. Someone who is indignant because a server refuses to seat them at a table when no tables are free is rightly blameworthy. Perhaps Wallace could say in this case that what matters is not the reactive attitude itself but the moral obligations that the agent accepts, as a matter of choice. But it seems to me that sometimes the reactive attitudes can reveal to us which obligations we accept, even if we never chose to endorse them. For example, I may begin to feel resentful in a relationship before I realize that it has become codependent and that I am committed to expecting my partner to carry more than their fair share. The reactive attitudes often arise prior to any explicit endorsement of moral obligations and can conflict with our considered views.<sup>37</sup>

Turn now to a different kind of choice-centered view that builds reasons and ends into the agent's choices. Call this view a maxim-type view. On this view, the agent is still held responsible for their conscious choices, but those choices represent an adoption of long-term intentions for action in the form of principles, e.g., "I will do x in y circumstance for z end." This view more readily accommodates responsibility for the agent as a reasoner, such that reason isn't merely some

---

<sup>37</sup> It seems to me that in these cases our reactive attitudes serve an epistemic function in helping us discover or pay attention to salient moral features. This helps to explain why some reactive attitudes can be inappropriate—sometimes they pick up on unrelated features or overly highlight some moral features of a situation while obscuring others.



precondition for responsibility—it's part of what we can be held responsible for. How might this kind of view accommodate for non-voluntary states?

Take Christine Korsgaard's (1996a) example of failing to consider her friend Charlotte—she fails to think of her friend when seeing something in a store that would be perfect for Charlotte, she does not worry when a catastrophe happens in Charlotte's neighborhood, and it never occurs to her to do anything to make Charlotte happy. Korsgaard concludes that:

Under these circumstances, surely Charlotte would be entitled to complain that there is no real sense in which I have her happiness as my end. It would not be pertinent for me to reply that I have no direct control over what occurs to me. To find certain features of the world salient is part of our notion of what it is to have an end (180).

Korsgaard's picture seems to broaden ends to include the more passive part of our nature. This can accommodate responsibility for facets of our moral lives that are not a direct result of our choices, which is a good-making feature of the view. The issue with Korsgaard's view is that adopting an end in the first place is “a volitional act,” or one that requires choice (180).<sup>38</sup> It is fundamentally a tracing account that boils down to choice as the core explanation for responsibility. There is no room for responsibility for the kind of unchosen dissent that Adams rightly picks out as a core part of who we are as moral reasoners.

## 2.4 Tracing Accounts

Tracing accounts often try to explain responsibility for non-voluntary reactions and feelings in terms of past choices.<sup>39</sup> The reaction is thus “traced” back to the choice that engendered it. For

---

<sup>38</sup> “To adopt an end is to perform an internal action. But it is also to undergo certain changes, changes in your representational capacities. It is to come to perceive the world in the way that having the end requires” (180).

<sup>39</sup> Tracing accounts can also function more generally and not simply with respect to choice. What is important is that X is traceable back to Y, where Y is something that the agent is clearly responsible for and where the relation between X and Y makes the agent responsible for X. For a prominent statement of a tracing approach, see Fischer and Ravizza (1998) (49-51). “The general approach we are developing is a ‘tracing’ approach: when an agent is morally responsible for an action that issues from a mechanism that is not appropriately reasons-responsive, we must be able to trace back along the history of the action to a point (*suitably related to the action*) where there was indeed an appropriately reasons-responsive mechanism” (50-51).

example, on this kind of account, I would be responsible for developing transphobic attitudes because I had made choices to follow anti-trans activists on twitter for their anti-trans content. The clearest tracing cases involve a direct connection between the initial choice and the resulting non-voluntary actions or feelings. A drunk driver is responsible for endangering others because they chose to drink heavily that night, even if at the time of the accident they were severely inebriated. Tracing accounts seem to get at least some cases right.

What is wrong with tracing accounts then? Why can't they get us what we want (namely control, avoidability, and fairness) and also recognize our moral practices that govern our feelings and attitudes? I think tracing accounts rightly capture that there is an important difference in our moral assessments if the attitudes and emotions in question are a direct result of the agent's choices.<sup>40</sup> An attitude that I have cultivated says something different about me than an attitude I have simply absorbed. It is unclear to me, though, that we should exclude the latter category from the set of things for which we are responsible. Our unchosen emotions and non-voluntary reactions can still appropriately elicit the moral reactive attitudes as they did in the case of the fatphobic friend at the pool. Prior choices matter if they are part of a pattern or otherwise bear on the present context, but you can also be morally responsible for non-voluntary attitudes that don't trace back to a set of clearly related choices. Tracing accounts don't *exclusively* get us responsibility for non-voluntary states.

As Smith (2005) argues, our judgments and reactions

do not always arise from conscious choices or decisions, and they need not be consciously recognized by the person who holds them. Indeed, these judgments are often things we discover about ourselves through our responses to questions or to situations. For example, I

---

<sup>40</sup> Smith (2015b) contends that volitional tracing accounts can't make good sense of responsibility for attitudes because attitudes are never under our direct voluntary control. This is because "these states are sensitive to our judgments about reasons, and we cannot generally control whether relevant reasons are present in the circumstances or not." Furthermore, "our attitudes, *by their very nature*, are judgment-sensitive states, so there is no need to try to 'trace back' to some *other* judgment-sensitive locus of responsibility to establish a link between the agent and the thing for which he is being morally criticised." For a more thorough discussion about whether some version of the control intuition can be made sense of with tracing accounts, see Vargas (2005) and Fischer and Tognazzini (2009).

may not realize, until I am faced with a choice, that I value the intellectual freedom and autonomy associated with a career in academia more highly than the economic rewards and benefits associated with a career in law (252).

Additionally, our attitudes and emotions often arise incidentally as a result of our choices, in a way that does not get us the sufficiently tight connection that tracing arguments require. Am I responsible for developing unhealthy attitudes about thinness as well as anti-fat bias *because* I have chosen to go out in the world, where I've incidentally encountered numerous billboards, advertisements, and other depictions of beauty as a particular kind of thinness? No. That choice does not plausibly explain responsibility for this unintended, unrelated, and unknown outcome.

I do think that we are responsible for such attitudes, though I cannot fully argue for that here. What I can say is that tracing accounts cannot get us responsibility for attitudes and non-voluntary reactions that are not connected to our choices in the right way. We should preserve responsibility both for our good non-voluntary judgments that point us in better directions as well as our less helpful non-voluntary judgments that point us in immoral directions.

## 2.5 Worries About Moralism

One might worry at this point that the kind of view I am pointing towards is too moralistic and will allow us to blame people for infinitesimal failings. Shouldn't it matter more that people choose some course of action or endorse something? Especially if they are working hard against some kind of internal proclivity that they recognize to be wrong?

Take Nagel's (1979) example of an envious person who chooses well:

Even if one controls the impulses, one still has the vice. An envious person hates the greater success of others. He can be morally condemned as envious even if he congratulates them cordially and does nothing to denigrate or spoil their success (32-33).

According to the moralism worry, it is too much to condemn people for traits that they do not endorse and do not act on. Envy is not something that is directly under our control, and changing that kind of character trait requires much time and effort.

One thing to say in defense of the view that I am carving out is that moral condemnation may be too strong a reaction to the kind of case that Nagel brings up, depending on the extent of the vice. We should not expect inhuman perfection of ourselves or of others. Why then, should we think that we are still responsible for our un-endorsed vices? Responsibility comes into play when those vices affect our relationships.

The mode of character is dependent on situational and practical reasons, but it can sometimes be reasonably applied in these cases. We sometimes rightly feel reactive emotions in the range of gratitude and resentment for the non-voluntary reactions of others. In Nagel's case, the vice is still relevantly part of the narrative of character; without it, we would not be able to fully understand the full import of the choices to do otherwise. What we want is an accurate yet charitable picture of ourselves that allows for constructive opportunities for growth and for making amends. Sometimes we need forgiveness; other times we need to simply count on our friends to not excessively dwell on all our flaws.

It is not helpful nor desirable to constantly ask in what minute ways we have failed to have the right attitudes or make all the right choices—interpersonal relationships do not flourish under this kind of scrutiny. It can be a vice to perseverate over one's actions and think over each thought and deed to determine how morally good, bad, or praiseworthy it was. And it can equally be a vice to do the same for the actions of others.<sup>41</sup> Sometimes we need to hold ourselves and others responsible for our attitudes that bleed out despite our best efforts or color what our choices actually represent,

---

<sup>41</sup> Matt King (2020) locates this vice as a problem of attending to the wrong things. Features of our relationships tend to govern the amount and quality of attention we should give to the wrongdoing of others.

but we need to do so with sensitivity and good judgment. To live well with ourselves and with others, we must be able to love and accept ourselves as imperfect and messy people, not merely as the set of things that we endorse and have direct control over. Part of that process is working to integrate the non-voluntary self and the voluntary self.

## Conclusion

In this chapter, I've gestured towards considerations that might point us away from a choice-centered view like Wallace's, though not from the insight that choice is important for understanding moral responsibility and character. In the next chapter, I turn to T.M. Scanlon's (1998) judgment-centered view of responsibility. Just as I have argued that a choice-centered view cannot get us everything we want in an account of moral responsibility, I will also argue that a judgment-centered view cannot get us everything we want. Both views are instructive for my own account, which attempts to combine insights from each into a synthesis view that can better capture our ideas about character.

### **Chapter 3: Scanlon's Judgment-Centered Account of Moral Responsibility**

In the previous chapter, I discussed R. Jay Wallace's (1994) choice-centered view of responsibility. In this chapter, I lay out T. M. Scanlon's (1998) judgment-centered view of responsibility. Together, Scanlon's and Wallace's views represent the starting point for my own project, which seeks to capture the kind of interplay between the voluntary, chosen aspects and non-voluntary, judgment-sensitive features of our agency. In this way, my project is also an outgrowth of Robert Adams' (1985) picture of the moral self which recognizes these different, mutually informing and contesting parts. Just as I have shown that Wallace's view cannot, without supplementation, get us responsibility for the reasons we hold nor for important non-voluntary reactions, I show that Scanlon's view must be expanded to capture responsibility for choices as something unique and distinct from non-voluntary, judgment-sensitive attitudes.

In Part 1, I lay out Scanlon's view of responsibility. In Part 2, I explain where I depart from Scanlon's view and why.

#### Part 1: Scanlon

In this section, I provide an overview of Scanlon's frame for thinking about responsibility and explain the role of judgment-sensitive attitudes. I then briefly discuss similarities and differences between Scanlon's and Wallace's approaches before turning to Part 2.

##### 1.1 Substantive Responsibility and Attributability Responsibility

Before I present Scanlon's judgment-centered view of responsibility, let's return to the distinction between substantive responsibility and accountability responsibility. It's important to revisit this distinction in Scanlon's own terms to understand his larger project as well as where he accounts for the value of choice.

For Scanlon, responsibility as attributability has to do with “questions about whether some action can be attributed to an agent in the way that is required in order for it to be a basis for moral appraisal” (248). When we ask whether someone is responsible in the attributability sense, “what we are asking is whether that person is properly subject to praise or blame for having acted in that way” (290). For example, someone might be responsible in the attributability sense if their action flows from one of their judgment-sensitive attitudes that is in principle under the control of their reason (272). But, if an action comes out of a blind urge or some other state that is not under the control of reason, that action is not attributable to the agent and they cannot be blamed or praised for it.

Substantive responsibility involves “substantive claims about what people are required (or, in this case, not required) to do for each other” (248). To say that someone is substantively responsible for some outcome is “to say that that person cannot complain of the burdens or obligations that result” (290). For example, imagine a city took all reasonable precautions to keep citizens away from a toxic part of the dump, and yet a resident managed to walk into that area and expose themselves to carcinogens. That resident could not reasonably complain about the burden they suffered, since the city took all reasonable precautions to keep the resident from entering that area. Or take someone who decides to enter the milk-selling business and who acquires certain obligations to sell safe milk as a result. The milk-seller cannot reasonably complain about these obligations being placed on them, as the occupation carries those obligations with it. Remember that substantive responsibility does not necessarily entail attributability responsibility; however, “insofar as punishment involves an assertion that the agent governed him- or herself in a way that was faulty, it is appropriate only when this is true” (267).

On the side of substantive responsibility, Scanlon recognizes the value of choice, though not in the same way as Wallace. Scanlon wants to understand why it is the case that “when a person could have avoided a certain result by choosing appropriately, this fact weakens her grounds for

rejecting a principle that would make her bear the burden of that result” (256). In other words, Scanlon sets out to explain 1) why principles that no one could reasonably reject are sensitive to individuals having the opportunity to choose, and 2) how considerations of responsibility can diminish the grounds from which one can reasonably reject a principle (251).

Scanlon concludes, somewhat surprisingly, that while choice has value to it, whether that be the value of self-determination, of being able to get the things we want, or being societally recognized as autonomous agents, if “a person has been placed in a sufficiently good position, this can make it the case that he or she has no valid complaint about what results, whether or not it is produced by his or her active choice” (258). On Scanlon’s view, what matters most in determining whether an obligation or burden is reasonable is not the fact of choice itself but the background conditions in which the agent finds themselves (260-1).

While I will critique Scanlon’s view of attributability for failing to recognize the importance of choice to assessing an agent’s character, Scanlon is able to capture the value of choice on the side of substantive responsibility, albeit through an understanding of the background values and conditions that make choice important in certain situations. In what follows, I will now focus on Scanlon’s view of attributability responsibility.

## 1.2 Scanlon’s Judgment-Centered View of Moral Responsibility

Scanlon’s overarching project in *What We Owe to Each Other* is to give a contractualist account of moral principles that govern how we relate to each other as rational beings. On this picture, legitimate moral principles are those that could not be reasonably rejected. This broad account has direct implications for Scanlon’s view of moral responsibility. First, “if an action is blameworthy, then the agent has either failed to take account of or knowingly acted contrary to a reason that should, according to any principles that no one could reasonably reject, have counted against his



action” (271). Second, since the principles that no one could reasonably reject are grounded in our value as rational beings, wrong actions flout “requirements flowing from another person’s standing as someone to whom justification is owed” (271).

To be a morally responsible agent, then, one must be a rational creature, because only rational creatures have the capacity of reflective self-governance. And the states that people are responsible for are their judgment-sensitive attitudes, which are “those attitudes that, in a rational creature, should be ‘under the control of reason’” (272).

Scanlon also endorses an answerability view of responsibility, such that we are responsible for those attitudes that “we can in principle be called on to defend... with reasons and to modify them if an appropriate defense cannot be provided” (272).<sup>42</sup> From this, it also follows that we are responsible for the range of things that are judgment-sensitive or that should be under the control of reason, namely our judgment-sensitive attitudes, as only they are appropriate targets of demands for us to provide justification or revise those judgments.<sup>43</sup>

For Scanlon, judgment-sensitive attitudes are the class of attitudes for which we can ask people to give reasons in the standard, normative sense, i.e., reasons that count in favor of something.<sup>44</sup> Judgment-sensitive attitudes include “beliefs, intentions, hopes, fears, and attitudes

---

<sup>42</sup> Angela Smith’s (2005) view of responsibility for attitudes similarly begins with a conception of responsibility as answerability but differs from Scanlon’s in a few ways. On her rational relations view, “an agent is responsible for  $\Phi$  just in case  $\Phi$  bears a rational connection to the agent’s evaluative judgments” (2012, p. 577). Note that the language here is that of evaluative judgments and rational connections, not of judgment-sensitivity. Smith argues that what we are ultimately concerned with is the evaluative content of the attitudes we hold and express in action: “When we praise or criticize someone for an attitude it seems we are responding to something about the content of that attitude and not to facts about its origin in a person’s prior voluntary choices, or to facts about its susceptibility to influence through a person’s future voluntary choices” (2005, p. 251).

<sup>43</sup> One might wonder, on Scanlon’s view, how an akratic agent can be held responsible for a bad action if all the judgments they hold are correct. The answer seems to lie in the answerability part of the view, because it still makes sense to demand an explanation from that agent, who would then report that they disavow the attitude. Expanding Scanlon’s notion of self-governance to include both the indirect control we have over some of our judgments and the direct control of choice could also help with these kinds of cases, because it better characterizes the conflict within the agent as one between judgment and choice, not as one merely within judgments.

<sup>44</sup> This kind of reason can be contrasted with reasons given simply for historical explanation or the reasons why we believe something. Scanlon is concerned with determining which reasons “really count in favor of the thing in question,” namely, which reasons can be taken to be good reasons (19). For a critique of Scanlon’s view of reasons in the standard normative sense as a unified class, see Ralph Wedgwood (2015). I suspect that the differences Wedgwood finds between

such as admiration, respect, contempt, and indignation” (20). These attitudes are “up to us” in the sense that “they depend on our judgment as to whether appropriate reasons are present” (22). They are “attitudes that an ideally rational person would come to have whenever that person judged there to be sufficient reasons for them and that would, in an ideally rational person, ‘extinguish’ when that person judged them not to be supported by reasons of the appropriate kind” (20). Rationality of the kind involved here (that an ideally rational person would have) has to do with the form of our judgments, namely the systematic connections between our judgments. This sense of rationality can be contrasted with the senses of “rationality” and “irrationality” that concern what reasons we ought to recognize or what judgments we ought to have (30).

While Scanlon is committed to something being in principle “under the control of reason” as a prerequisite for responsibility, he specifies that responsibility for judgment-sensitive attitudes does not require choice or decision, nor does it require conscious judgment (21-22).<sup>45</sup>

A person can be criticized, and asked to provide justification or acknowledgment and apology, for things that seem to have been done inadvertently in a situation in which advertence is called for. Being in principle “under the control of reason,” and arising from conscious judgments or choice, are two different things (272).

If neither consciousness nor choice is required for responsibility, then what states will be excluded?

States for which it does not make sense to demand normative reasons, such as mere feelings like

---

the normative-explanation role of reasons and the ideal-motivation role of reasons (among other types of reasons) should be explained by reference to different practical standpoints and not at the level of reasons themselves. Wedgwood’s argument at most seems to get us the claim that reasons and the practical standpoints from which different reasons are salient are so closely tied together that it is unclear one can be considered more fundamental than the other. Mark Schroeder (2007) rightly picks out this relation between reasons and practical standpoints and claims that the basic normative reason relation is a “three-place relation between the thing that is the reason, the agent for whom it is a reason, and the action-type that it is a reason to perform” (17). I differ from Schroeder in that it is unclear to me that reasons always count in favor of actions (it seems that reasons can count in favor of having a certain attitude, etc.).<sup>45</sup> It also does not require that the agent’s judgment-sensitive attitudes are modifiable (276). “The idea that moral criticism offers reason in the light of which an agent is asked to reassess his or her judgment-sensitive attitudes does not in fact lead to the conclusion that being unable to appreciate the force of a moral reason makes an agent immune from moral criticism for failing to give that reason appropriate weight in deciding what to do.... If a person holds a certain judgment-sensitive attitude, then, because this attitude is in principle sensitive to and dependent upon his judgment, it is appropriate in a general sense to ask him to defend it or to disown it” (289).

hunger or tiredness, as well as the class of blind urges, as those urges do not involve tendencies to see something as a reason for acting (20, 273).<sup>46</sup>

What about irrational attitudes that seem to be impervious to our considered judgments? Or, what happens when “there is a direct clash between the judgments a person makes and the judgments required by the attitudes he or she holds” (25)?<sup>47</sup> Return to Thomas Nagel’s (1979) case in which someone’s attitudes and character are, on the whole, quite bad and yet he manages to act perfectly on the basis of his considered judgments:

A person may be greedy, envious, cowardly, cold, ungenerous, unkind, vain, or conceited, but behave perfectly by a monumental effort of will. To possess these vices is to be unable to help having certain feelings under certain circumstances, and to have strong spontaneous impulses to act badly. Even if one controls the impulses, one still has the vice (32-33, quoted directly by Scanlon on 273).<sup>48</sup>

Assume that this person overrides his impulses because he holds correct moral judgments and chooses to act in line with those judgments.

Scanlon argues that the spontaneous impulses the agent experiences are “not blind urges but tendencies to see certain considerations as reasons for acting in certain ways” (273). Furthermore, the person who has these tendencies “must see these tendencies as *his*: as tendencies to take certain judgment-dependent attitudes that he has to overrule and correct” (273). For Scanlon, these cases are not conflicts in which rational capacities “are overmastered by some other force,” but instead

---

<sup>46</sup> Of course, hunger, boredom, and tiredness can often reflect other evaluative judgments. For example, I may become hungry and tired in a class because I am uninterested in the material and wish to leave. Even if I am not responsible for the hunger itself, I may still be responsible for the lack of interest or other attitude that is informing my physical state.

<sup>47</sup> Shoemaker (2011) argues that in cases of irrationality, in addition to the presence of two conflicting attitudes, there is the fact that those attitudes present a conflict within the agent. There is something about the form of the relation between the two attitudes that allows us to identify this as a case of irrationality. Contrary to Shoemaker, Smith (2012) argues that the agent is responsible for her irrationality simply as a consequence of her responsibility for each of the conflicting attitudes she holds. This is because Smith thinks that the kinds of demands we put to others concern the judgments they hold, and so the problem in cases of irrationality is not the conflict itself but that one of the two judgments is not correct. I think that Smith is right to say that we care about the judgments that others hold, but I think it’s plausible that we also care about how well someone is integrated as a moral person. It may be reasonable in some cases to levy the charge “why can’t you see that your views are always so inconsistent!”

<sup>48</sup> As Scanlon notes, “Nagel’s examples thus have the same structure... as an example of a common form of irrationality, of a person who keeps thinking of the approval of a certain group of people as something to be sought even though he in fact judges this not to be a reason, and always dismisses it when it occurs” (273).

“conflicts within a person’s rational capacities” (273-274). On Scanlon’s account, then, the irrational attitudes (both the judgments the agent tries to overrule and the judgments the agent chooses to act on) are attributable to the agent for the purposes of moral assessment, as it is “appropriate to ask him to defend or modify and retract them” (274). He concludes the answer to determining what kind of moral criticism is appropriate will be “a boringly mixed one: the agents have governed themselves well in dealing with these tendencies, but they would be better people if they did not have them” (274).

What does this self-governance require? For Scanlon, “[a] person governs herself in the sense required if she is sensitive to the force of reasons and to the distinctions and relations between them and if her response to these reasons generally determines her subsequent attitudes and actions” (281). If either of these conditions fails to be met, then the person “cannot be a participant in a system of co-deliberation, and must be seen, rather, as simply a force to be dealt with, like an animal” (280). The standard for moral criticism is thus that an agent “has governed herself in a way that would not be allowed by any principles that no one could reasonably reject” (268).<sup>49</sup>

### 1.3 Scanlon vs. Wallace

Note that these dual conditions for self-governance (the ability to recognize reasons and the connection between reasons and actions) very closely mirror Wallace’s prerequisites for responsibility, which he calls the “powers of reflective self-control: the general ability to grasp and apply moral reasons and to regulate their behavior by the light of such reasons” (154-155). Likewise,

---

<sup>49</sup> Interestingly, Scanlon does not claim that the ability to understand moral reasons is required to be morally responsible. Scanlon argues that this ability is unnecessary, as “a person who is unable to see why the fact that his action would injure me should count against it still holds that this doesn’t count against it” (288). Even if the agent is “unable to see the force of some reason that counts against this attitude, this does not alter the fact that the attitude and the judgment that it is warranted are properly attributable to him. Any errors involved in these attitudes are also attributable to him, and he is therefore properly criticized for holding them” (289).

Scanlon follows the general move from our practices of holding responsible (answerability) to the states and capacities required for responsibility. Scanlon and Wallace diverge in their emphasis on which ability is primary for responsibility. For Wallace, choice makes the agent responsible; reason is a precondition. For Scanlon, judgment-sensitive attitudes make the agent responsible in the attributability sense; choices matter insofar as they express attitudes.

In the previous chapter, I argued that Wallace's choice-centered view couldn't capture the importance of judgment for moral responsibility. In the rest of this chapter, I show that Scanlon's judgment-sensitive view cannot account for the importance of choice to attributability.

## Part 2: Departures from Scanlon

Scanlon's view is perhaps closer to mine than Wallace's. The judgment-centered view of responsibility captures responsibility for the reasons we recognize and act upon, and it also captures conflicts internal to the agent between different judgments the agent holds. There are two main departures I take from this kind of view: 1) Without supplementation, there is not a sufficiently clear distinction between the kind of direct influence we have over our choices and reflective judgments and the kind of indirect influence we have over other, more stubborn and automatic judgments. 2) Likewise, without an additional story about the agent's history and context in which they acquired certain attitudes, there can be too great a focus on the content of the attitudes themselves and not on how well the agent reasoned given the information that was available to them. There are resources in Scanlon's judgment-centered view that point towards how we could start developing both of these additions, though my account is not merely an extension of Scanlon's—there are lessons that I draw from both Scanlon and Wallace, but my positive account more strongly emphasizes history and the importance of narrative.

## 2.1 Direct and Indirect Control

To see how Scanlon's judgment-centered view might have initial difficulty in handling the distinction between direct control and indirect control, let's look at Scanlon's notion of self-governance. As Wallace argues, there is a morally relevant distinction between things that are under our direct influence, like our choices, and things that are under our indirect influence, like our beliefs. Unlike Wallace, I do not want to claim that direct control matters more for responsibility. I simply want to claim that the distinction between direct control and indirect control matters for narratives of character and how we interpret them. Scanlon's notion of self-governance appears to run these together.

Return to Nagel's case in which an agent has exceedingly bad attitudes but manages to behave in a way that is morally appropriate. On Scanlon's account, the agent's attitudes are attributable to the agent for the purposes of moral assessment, so long as it is "appropriate to ask him to defend or modify and retract them" (274). According to Scanlon,

the relevant question of moral assessment in such a case concerns the kind of criticism that is appropriate, taking into account both the attributability of these attitudes to the agent and his attitude toward them. The answer is a boringly mixed one: the agents have governed themselves well in dealing with these tendencies, but they would be better people if they did not have them (274).

I agree with Scanlon that the answer is a mixed one. The interesting question is how Scanlon's view can get at the distinction between what comes naturally to the agent (the agent's bad attitudes that are under the agent's indirect control) and what the agent chooses to do (the agent's morally good actions that are under the agent's direct control).

When analyzing Nagel's case, Scanlon seems to recognize the difference between directly overriding/overruling/affirming judgments and merely having the tendency that must be indirectly controlled:

On the one hand, it seems not to be: the greedy person he describes judges correctly that it would be wrong to seek his own advantage in the ways in question, and he governs himself

accordingly. On the other hand, the tendency to think that seeking his advantage is what he has most reason to do is also attributable to him, even though he overrules this tendency when it occurs, and affirms the opposite judgment.... The greedy person must see these tendencies as *his*: as tendencies to take certain judgment-dependent attitudes that he has to overrule and correct (273).

It's hard to even construe this case without thinking about the agent's choices as being opposed to the agent's judgments. If we try to think about this case as simply a conflict between judgments, it becomes difficult to see how the agent is best described as greedy—if he is sufficiently internally divided and his correct judgments always win out, is he just greedy or is he also remarkably circumspect and self-controlled? On Scanlon's view of self-governance, “a person governs herself in the sense required if she is sensitive to the force of reasons and to the distinctions and relations between them and if her response to these reasons generally determines her subsequent attitudes and actions” (281).

Because Scanlon places everything at the level of reasons-recognition and reasons-responsiveness, we lose, at the theoretical level, important features of our internal landscape that represent quite different internal states. As Robert Adams (1985) argues, it can be helpful to have a more complicated picture of self-governance that has a division of powers by which the non-voluntary can push back against the voluntary, and vice versa. This distinction is necessary to understand conflicts within the agent between the attitudes that the agent has rejected (that they can only indirectly control and adjust) and the attitudes the agent has chosen and endorsed (that they have direct influence over). Scanlon's account of self-governance seems to cover both under the broader category of judgmental activity, and his account could be expanded to allow for this interplay. Scanlon implicitly recognizes the distinction between direct and indirect control when analyzing Nagel's case but does not strongly build it into his definition of self-governance.

When constructing narratives of character, it matters to us what the agent actually chose to do as well as what underlying judgments informed the agent's choice. It seems that in order to get a

full picture of who we are as moral agents, we must recognize the importance of choice. This more direct part of self-governance is not merely an expression of the set of judgment-sensitive attitudes that we already hold; it is an executive decision about which of our attitudes to endorse in action. If we look to the agent's choices, we can see which attitudes he is working to cultivate, which ones he continually chooses to act upon, and which he is seeking to eliminate. This is not a choice from nowhere; it is a choice within the reasons and attitudes the agent has.

## 2.2 Varieties of Indirect Control

Additionally, because the focus is so narrowed to judgment and judgment-sensitivity, Scanlon's view and views like it tend to problematically collapse different kinds of judgment that may be more or less reasons-responsive into the same category. Take the case of an arachnophobe who judges that spiders are relatively harmless yet remains deathly afraid of them.<sup>50</sup> According to the judgment-sensitive view of responsibility, it appears in this case of irrationality that both attitudes are attributable to the agent. At best, it seems that this agent can indirectly control the arachnophobic thoughts, as they are not extinguished when they normally should be. In this case, the agent has no direct control over whether the rational connections between their judgments and attitudes hold, but they do have direct control over taking actions that should help to change their arachnophobic attitudes to fit with their considered judgments. The agent may, for example, choose to try to extinguish their fear of spiders by exposing themselves to terrifyingly accurate plastic spider figurines. But note that at this point we're operating at the level of brute conditioning and not at the level of thinking about reasons that count in favor of or against the arachnophobia. While we all likely have to engage in some form of conditioning or habituation to change features of ourselves, some judgments are much more sensitive to reasons than others.

---

<sup>50</sup> For another discussion of the arachnophobe and irrationality, see Smith (2008, pp. 253-255).



The lack of control that the agent has over their arachnophobic judgments does not necessarily carve those judgments off from the realm of moral responsibility if those judgments affect their relationships with other people. However, the fact that they only have attenuated indirect control at best over those judgments puts them in a different class of judgments than the ones that they can more freely change and that are more responsive to reason. In positive cases, stubborn judgments about the worth of human life can be laudable, but in negative cases in which the agent is internally conflicted and wishes to jettison stubborn false beliefs, the lack of control often modulates our actual expectations and reactions to those individuals. The kind and degree of change we expect may vary. The mere fact of arachnophobic attitudes does not tell us much. The recalcitrance of the arachnophobic attitudes + the agent's reflective judgments + the agent's chosen courses of action tell us much more. The agent may be doing everything we could reasonably expect, even if the irrationality is still relevant to telling a full narrative of character.

### 2.3 Embodiment

There is another point to be made about judgment-centered accounts (and most contemporary accounts of moral responsibility, for that matter), which is that they often present this picture of human beings as almost ethereal or disembodied rational creatures operating at the level of reasons or choices alone. Scanlon's view sidesteps some of these worries, as he recognizes the kinds of internal conflicts and inconsistencies that we often have as messy human beings. Where it falls short is in the exclusion of certain bodily states. We are embodied creatures, and it seems important to understand how those features of our experience line up with our judgment-sensitive attitudes and subsequently who we are as moral agents. We are not pure reasoners or ideal mental agents dissociated from any physical limitations or contexts.

Take the part of Scanlon's view that separates out states such as hunger or tiredness as states that we cannot be responsible for, because these feelings and other blind urges do not involve tendencies to see something as a reason for acting (20, 273). It is often the case, however, that our bodily feelings are directly related to various reasons we recognize. As I am writing this, I am still finding myself feeling tired, with a knot in my stomach, and with other physiological dysregulation as a result of (or perhaps even constituted by) my rational response to distressing political events that may impact my ability to live the good life. These bodily states often serve as signs and signals that there is something important to be paid attention to (and, in other cases, they may make it much harder for us to pay attention to other things or see things in a different light). Even if there is some broader category of non-human moral persons that we are a part of, there are specific moral experiences and expectations that we have as human beings, and we should also pay attention to those features of who we are as human persons in order to give a view of morality that makes sense within our own lives. It seems like a Scanlonian view could accommodate this, so long as it recognized a deeper relation between our reason and our embodied states.

#### 2.4 Historical Context

A final place of departure from judgment-centered views is from their narrow focus on the content of and relations between the attitudes and judgments themselves, which separates out the context that informed them. We can see this by looking at both Scanlon's judgment-centered account as well as Angela Smith's (2005) account which builds on Scanlon's. Smith's account claims that what we care about for the purposes of moral assessment are the evaluative judgments we hold that are rationally connected to our attitudes. What happens, though, if two people come to hold the same evaluative judgments or attitudes from very different historical contexts?

Take a case in which there are two people, both who hold identical homophobic attitudes. The first grew up in a homophobic enclave with no outside information. The second had a good upbringing but sought out homophobic organizations and anti-gay arguments in adulthood. Intuitively, it seems that our judgments of blameworthiness might differ in this case, but what might explain that difference, given that the evaluative judgments are the same? Smith's answer to this case is that "this question of responsibility (namely, the responsibility one has for becoming a certain kind of person) must be distinguished from the question of one's responsibility for the attitudes one in fact holds" (267). In other words, Smith's account can tell us that both people are responsible for their homophobic attitudes, in the sense that those attitudes are attributable to them, but that this should be separated from any questions as to whether they are responsible for becoming the kind of person that they are. The historical question, or context in which the attitudes were formed, is separated from the content of the attitudes that the agent currently holds.

Like Smith, Scanlon wants to separate out the question as to "whether it is appropriate to take [an agent's] actions as indicating faulty self-governance" from the question as to whether "he has negligently allowed himself to fall into habits and associations that have undermined his character" (284-285). This kind of stance makes sense in certain contexts. For instance, it is sometimes the case that someone's background or history will be used to excuse them of wrongdoing or distract from the wrong that has been committed. That is something to be avoided. Additionally, we do not always need to determine the historical context, and it may be more useful in some cases to solely address the content of the judgment at hand. At the same time, I think it is important to be able to make more fine-grained moral judgments that can make sense of the information that people had available to them and how well they reasoned and chose within their circumstances. I think it's possible to do this without separating out the historical question.

If we build in context into our assessment of the agent's attitudes, we can see why there might be a moral difference between the two homophobes. At one point in time, I was the first homophobe. I think it's plausible that I reasoned at least not horrifically poorly given the information available to me within a relatively sheltered environment, but I still hold myself responsible for my past homophobia. At the same time, I changed my views once I was able to give myself time to process what that would mean for my place within my family and church community. The second homophobe, however, seems to have clearly reasoned incredibly poorly. They had access to information that I did not and still chose to endorse their homophobic attitudes and inculcate them within themselves. They should have known better.

What emerges from these two cases is the idea that it is not always the content of the judgments or attitudes themselves that should be our only focus but also how well we reasoned given the information available to us. The context sometimes changes the moral assessment. We never make our judgments within ideal circumstances—all of our attitudes and judgments are context-dependent to some degree. We aren't interchangeable rational beings with propositional sets of attitudes, we're human beings who live in particular communities and have certain histories.

## 2.5 Inferring Judgments

This last section is less a departure and more a note about how judgment-centered views can sometimes wrongly appear to give us a clear picture of the agent's internal states based solely on inferences from their behavior. One tricky piece of Smith's view is that her examples sometimes require us to mind-read judgments from the behaviors someone is performing, as, for example, from forgetting a friend's birthday to the judgment that she must have begun to care less about her friend. While these inferences may work for some, they can be very tricky for neurodivergent people. As

August Gorman notes in the abstract for their book [\*Altering the Fault Lines: How Neurodiversity Shapes\*](#)

[\*Moral Responsibility:\*](#)

According to a prominent tradition, moral responsibility is constituted by our proneness to hardwired reactive attitudes that assume that our relationships have a measure-for-measure reciprocal nature. It is important that we are constantly monitoring each other subconsciously through emotional ‘mindreading,’ and the ubiquitous automatic responses we have to one another’s attitudes are an essential component of participatory interpersonal relationships. I argue that this is false, and point to the fact that autistic communities navigate rich interpersonal relationships among themselves without engaging in this sort of monitoring.

Neurodivergent people may be baffled by the mindreading they are expected to perform and find themselves to be the targets of reactive attitudes that are unwarranted because a neurotypical person has assumed that a certain intention or mental state was there that was not in fact there. This kind of false accusation due to bad assumptions can be quite harmful. There are resources in the answerability view of responsibility to ask individuals to explain and justify themselves instead of assuming, but asking others to continually justify themselves when it is not in fact warranted due to ongoing misunderstandings may still constitute a moral harm.

Elizabeth Spelman (1978) likewise notes that this kind of moral mind reading is insufficient for understanding other people:

One sometimes gets the impression, reading the philosophical literature on the problem of other minds, that it really does not matter what you say about yourself, since others could come to have knowledge about you independently of what you say. But this would only be true if the details of your mental life could be portrayed, or betrayed, in the details of your behavior. And they cannot. The body is too crude an instrument for such portraiture; it muffles the message (154).

Add to this the possibility of competing explanations of behavior that move between the medical and the moral mode, and things get even more complicated, as when my depression from internalizing bad theology was explained away as merely a result of my hormones.

Because it can be difficult to infer attitudes from behavior, we will need to rely, in part, on the testimony of the person we are holding responsible to determine whether our initial judgment

was accurate. Spelman argues that to treat someone as a person, we need to respond to the person that they are, which is the person someone takes themselves to be:

We can mark the two different senses of treating someone as a person in this way: it is one thing to treat someone as a person, as a bearer of rights, and another thing to treat someone as the person one is. Under the first rubric, in treating someone as a person I respond to his or her status as someone who has rights. Under the second rubric, I respond to the person someone is or, more exactly, the person whom someone takes himself or herself to be (151).

I think Spelman is on the right track in claiming that responding to the person someone is involves the self-conception of that person; however, it also requires an assessment of how that person *actually* is. Blatant hypocrisy, willful self-ignorance, and obliviousness to who one actually is should make a difference as to how we respond to those individuals. Whether there is a disconnect between a person's self-conception and who they actually are or the two are perfectly aligned, both the self-conception of the person and their actual set of attitudes and judgments may bear on their character.<sup>51</sup> I do think, however, that Spelman is exactly right to say that we should be responding to the person as a particular individual.

While judgment-centered views of responsibility may appear eminently clear in inference and application at the outset, things can very quickly get complicated.

## Conclusion

If choice-centered views fall too much on the side of emphasizing the voluntary and ignoring the non-voluntary, then judgment-centered views fall too much on the side of emphasizing the non-voluntary and ignoring the voluntary. In order to get a full picture of moral responsibility,

---

<sup>51</sup> Note that if something is contained in one's self-conception it is not necessarily endorsed, and we may hold multiple competing self-conceptions. Self-conceptions can be quite complicated: "they typically not only involve descriptions you actually satisfy, but also incorporate descriptions of the person you would like to be. The details of our hopes about what we might become are not worn on our sleeves; and even if others can make close guesses about our projects, they are not likely to know how large or small a role such projects play in our pictures of ourselves without asking us. Moreover, we change, and our conceptions of ourselves change; where such changes are not gross, other people are at a loss to register them without hearing from us what we take the changes to be" (Spelman, 155).

we need to capture both. Both judgment-centered views and choice-centered views are wrong to identify the conditions of responsibility with judgment-sensitivity alone or with choice alone. Both aspects of who we are as moral persons are needed to make sense of complicated internal conflicts in which our judgments and choices come apart. What we need is an account that can also build into it the embodied and historical contexts in which we make choices and form judgments. In the next chapter, I address how history and context should (and shouldn't) inform a view of moral responsibility.

## Chapter 4: How History and Context Matter

How much does history matter for responsibility? How does it potentially shape or mitigate responsibility?

On Susan Wolf's (1990) view, traumatic upbringings and other histories that impede our ability to recognize what is right and true can exculpate or remove responsibility. Becoming a worse person can thus mitigate responsibility. Along different lines, Fischer and Ravizza (1998) claim that an agent must have the right historical connection to, say, their actions in order for those actions to be the agent's own. The agent must take responsibility at some point in order to be responsible. Both views conceive of history as something that turns the dial of responsibility on or off. I argue that these views are too quick to dismiss responsibility due to a poor upbringing or a failure to take responsibility. Contra the views considered in this chapter, I will think of history primarily as something that helps to identify patterns of behavior and that provides crucial context to some narratives of character.

### Part 1: Wolf's Reason View

Susan Wolf's Reason view of moral responsibility does not explicitly build history into the conditions for responsibility. However, Wolf's claim that responsibility requires flexibility in thinking and the ability to get things right rules out responsibility in a number of cases in which agents have a poor upbringing or live in a morally abhorrent historical context. I argue that this is the wrong conclusion to make in these kinds of cases and that we should reject Wolf's Reason view.

#### 1.1 Wolf's Reason View

According to Wolf (1990), the central "ability that is crucial to responsibility is in fact the ability to act in accordance with Reason" (68). The ability to act in accordance with Reason requires



that the agent “is able to form her actions on the basis of her values *and* she is able to form her values on the basis of what is True and Good” (75).<sup>52</sup> To determine the responsibility of bad-acting agents, “we need to know whether she could have known better and whether, knowing better, she could have acted better” (88).

Wolf claims that the ability to act in accordance with the True and the Good requires that one “be sensitive and responsive to relevant changes in one’s situation and environment—that is, to be flexible” (69). These claims together entail that people whose values are shaped by traumatic upbringings in ways that make them less likely to see the True and the Good are less responsible, and that people whose reasoning is inflexible due to mental illness are also less responsible. Insofar as one has some inability to see the True and the Good or change one’s values in accordance with it, whether due to upbringing or mental constitution, one is less responsible and perhaps not even responsible at all, depending on the severity of the inability (87).

In her more recent work, Wolf expands her account of the necessary powers of reason an agent must possess in order to be responsible. To be a responsible agent, one must have an “active intelligence” which is understood to include a variety of faculties including imagination, perception, reason, empathy, etc. (2015a, 2015b).<sup>53</sup> The responsible agent must have “similar powers of perception, reason, and imagination that we do, so that when she responds to the world, it is the same world as ours to which she responds” (2015a, 140). Wolf argues that these powers are faculties

---

<sup>52</sup> I am using Wolf’s capitalization. This is a refinement of the sane deep-self view that Wolf puts forth in “Sanity and the Metaphysics of Responsibility” (1987/2003). The sane deep-self view requires “that a responsible agent be able to govern her (or his) actions by her desires and to govern her desires by her deep self” and furthermore that that deep self be sane, that is, that the agent has “the minimally sufficient ability cognitively and normatively to recognize and appreciate the world for what it is” (381-2).

<sup>53</sup> In these later iterations of the reason view (mainly, see “Responsibility, Moral and Otherwise”), Wolf (2015a) is concerned with giving conditions for “deep” responsibility more generally, not just moral responsibility, which Wolf takes to be but one kind of “deep” responsibility. Wolf thinks that we are sometimes deeply responsible for aesthetic works that we produce, for our senses of humor, and for other products of our active intelligence.

that we exercise, faculties of reason or intelligence, which give “us a basis for thinking that we may be not just agents, but *responsible* agents in particular” (2015b, 370).

So, on Wolf’s view, agents are responsible for actions that they could have performed in accordance with Reason, which requires that they are flexible in their reasoning, able to acquire the right values, and can govern their behavior accordingly.<sup>54</sup> In what follows, I challenge Wolf’s view based on three types of historical cases Wolf discusses (that are not mutually exclusive): the case of mental illness, the case of a poor upbringing, and the case of hardening one’s own heart.

## 1.2 Mental Illness

Take the first kind of case, that of mental illness. Wolf (2015b) is committed to the claim that we are only responsible for those acts which flow from our character and that certain mental illnesses are not part of our characters due to their lack of flexibility and context-sensitivity. According to Wolf, one must have an active intelligence that is flexible and context-sensitive in order to have the ability to form one’s values on the basis of the True and the Good. And, the ability to form one’s values on the basis of the True and Good is required for responsibility.

For example, Wolf claims that the vice of dishonesty should be separated from a pathological urge to lie (367). Since pathological lying is not flexible and hence does not flow from

---

<sup>54</sup> Note that in the original statement of the view, the focus is on the agent’s ability to act in accordance with Reason. While Wolf emphasizes Reason more than choice and requires that the agent is able to reason correctly, there is a question as to whether her view comes out to be roughly similar to Wallace’s (1994) view, which offers a similar set of abilities necessary for responsibility: “the general ability to grasp and apply moral reasons and to regulate their behavior by the light of such reasons” (154-155). Does Wolf’s view allow for responsibility for attitudes?

The best candidates for a pro-responsibility-for-attitudes reading are found in “Responsibility, Moral, and Otherwise” (2015a) when Wolf claims that a person might be responsible for her “particular features... or her actions” or her “actions and responses that manifest different aspects of [her] individual character” (140,141). Whether those “particular features” or “responses” include attitudes is somewhat unclear. It is also telling that the emphasis on action and behavioral control falls away in “Character and Responsibility” (2015b) and that the focus remains on developing the idea of an active intelligence or reasoning ability.

I think it’s plausible that at least some version of Wolf’s view might be able to accommodate responsibility for both choices and judgment-sensitive attitudes. The focus of this section is on a different claim that Wolf makes: that you need to be able to form your values on the basis of the True and the Good.

one's active intelligence, it is not a vice (and subsequently not an aspect of character). On Wolf's view, both vices and virtues—which are within the realm of the responsible self—require context sensitivity. However, it seems that many vices display a lack of context-sensitivity and flexibility and, more strongly, that some might *be* vices in part because of their inflexibility.

Think about the kind of legalism and fundamentalism that are exhibited by certain Christian circles. If ever single mistake merits eternal damnation and the rules must be followed exactly, then believers are left with a rigid structure that allows no deviation. This punishment-focused, rule-based structure lends to individual vices that are characterized in large part by their inflexibility. It's not merely that they've gotten the rules wrong; something has gone wrong at the level of focusing so intently on rules applied without context sensitivity. For example, when my classmate at private Christian school accidentally walked through the wrong door during school pickup (which was only mildly disruptive to the process), she was required to write an apology letter and face punishment as if she had done so intentionally. Those were the rules! The teachers that most strongly internalized this system were by far the most insufferable. Resentment seems to be appropriate in this case.

As another example, take an oblivious person. By virtue of being oblivious to their situations, they consistently fail to act in context-dependent ways. And yet, their obliviousness may still reveal certain aspects of their character. Perhaps they are so self-absorbed that they fail to care about anything else other than their own thoughts. Perhaps they don't even notice how their actions affect others because they've been consistently coddled by those around them. Think about the white male CEO who has never been challenged in his life, who tramples on the people around him, and whose underlings resent him for this ongoing pattern of behavior. Obliviousness thus appears to be a ripe candidate for a vice, even though the oblivious person lacks some kinds of context-sensitivity. Even if the obliviousness develops into a pathology, it is unclear that the frame of responsibility is inapt.

Kate Abramson (2016) convincingly argues that contemporary views like Wolf's which attempt to separate character traits from mental illnesses all fail to give a tenable account that can firmly separate the virtues and vices from medicalized diagnoses. As the previous two examples of legalism and obliviousness show, Wolf's inflexibility criterion cannot capture the fact that the medical mode and the mode of character sometimes overlap. Think about, for example, someone who becomes depressed after the death of a loved one. Though it is right to classify their response as depression, the person's grief is also an apt response to the situation and demonstrates the care and love they felt for the deceased. As such, it implicates the agent's character. Even if Wolf's claim that one must be able to form one's values on the basis of the True and the Good holds, the flexibility condition for moral responsibility must be dropped.

### 1.3 Poor Upbringing

Take the second kind of case, namely that of a poor upbringing. In this kind of case, we might imagine that an agent's childhood was so traumatic that once they reached adulthood, they could not recognize the True and Good or form their values in line with it. Take Wolf's (1987/2003) case of JoJo as an example. JoJo is the son of a dictator and is brought up to torture others and condemn them to death. Furthermore, JoJo values these activities and sees himself as following in the footsteps of his father (379).<sup>55</sup> Since JoJo presumably cannot see the True and the Good because

---

<sup>55</sup> For another fictional version of JoJo who is more fleshed out as a character, think about Azula in the original *Avatar: The Last Airbender* series. Azula is the daughter of Firelord Ozai, a genocidal, empire-driven leader of the Fire Nation. Azula enthusiastically takes up her father's aims after being raised to further the Fire Nation's expansion projects and commits a number of atrocities while seeking her father's approval. In the end, Azula still cannot gain her father's recognition and loses her sanity because of it. She is widely considered to be both a victim and a villain.

Watson (1987/1993) discusses a similar, real-life case of Robert Harris, who as both a victim and a victimizer, evokes conflicting responses in us. "The sympathy toward the boy he was is at odds with outrage toward the man he is.... In fact, each of these responses is appropriate, but taken together they do not enable us to respond overall in a coherent way" (138). Watson thinks that we find these cases troubling in part because of "the thought that if I had been subjected to such circumstances, I might well have become as vile. What is unsettling is the thought that one's moral self is such a fragile thing" (139).

his sense of morality has been so warped by his early experiences, Wolf claims that “it is dubious at best that he should be regarded as responsible for what he does” (380). Wolf thinks that it is unlikely that anyone with an upbringing like JoJo’s could have turned out differently, and that this too tells against JoJo’s responsibility (380). And yet, while I agree with Wolf that JoJo has been strongly shaped by his upbringing, it is unclear to me that he lacks responsibility.

For example, as a child, JoJo might have been horrified by the things his father asked him to do. Perhaps he was punished severely enough for this response that he learned to mask it and even came to somewhat enjoy these activities because of his father’s praise. It still might be possible, in adulthood, for JoJo to rediscover the feelings he had as a small child and begin to work to reform his value system in accordance with what is True and Good. Granted, that challenge will be a much greater one for him than for a child who had a good moral education, but it is unclear that JoJo lacks the ability to form his values on the basis of the True and the Good merely on the basis of his upbringing.<sup>56</sup>

Just as Michele Moody-Adams (1994) argues that agents are not determined by culture and that they themselves shape and uphold the cultures they live in, so too we might think that children are not entirely determined by their upbringing and can themselves contest or uphold certain features of what they were taught when they reach adulthood. Think about the common trope of the “rebellious teenager” or the many adults who hold somewhat radically different views from their parents on core moral issues.

We might worry that Wolf’s description of the case of JoJo has some similarities to how black children growing up in poor, violent neighborhoods have been portrayed as unable to avoid

---

<sup>56</sup> Counter to Azula, Zuko (the prince of the Fire Nation and Azula’s brother) experiences a similar upbringing but exhibits less cruelty at a young age and even stands up against his father, for which he is disgraced. Because of his uncle’s influence in his adolescence and despite a number of backsteps and mistakes, Zuko eventually redeems himself. While Zuko is in better circumstances than Azula, Zuko and the other characters around Azula are eventually able to see when she has gone too far. Azula’s loyal companions also leave her in the end, even though they’ve had even less exposure to other perspectives than Zuko has.

becoming criminals themselves due to the psychological harm of their upbringing.<sup>57</sup> If Wolf's conclusion in the case of JoJo holds in the case of these children, then they lack responsibility. While our circumstances are relevant to our assessments of character, we should be very careful about denying others the status of responsible agents. If one is not a morally responsible agent, then one is likely outside the moral community or only included as a ward of the moral community. And, for oppressed groups who have traditionally been viewed as less than human, the conclusion that many of the members of those groups are not responsible agents is a deeply worrisome one.

I think that a better diagnosis of the difference in our judgments between the case of someone with a good moral upbringing and the case someone with a traumatic childhood is that we understand that the latter person is also a victim. We may understand that it takes more work for the person with a traumatic childhood to adjust their behavior and beliefs, and we may recognize that they have some claims on us as a community to aid them in their recovery. However, this doesn't mean that we can't expect them to learn and change.<sup>58</sup> I am, of course, talking about our moral expectations here and not merely our predictions of how others will think and behave. More importantly, if we discover that an agent's circumstances were poor in some way, we are often operating at the level of excuses and not exemptions—we're giving more content to their attitudes and decisions, not necessarily employing the objective stance.<sup>59</sup>

As we have seen in both the case of mental illness and in the case of a poor upbringing, Wolf's view of the morally responsible agent fails to grant responsibility where it ought. Just because something is classified as a mental illness does not mean that we cannot be appraised on the basis of it. Just because someone was raised to have certain values, that does not mean that they cannot

---

<sup>57</sup> Moody-Adams (1994) discusses these kinds of cases as well (304).

<sup>58</sup> The kind and degree of change we expect may differ drastically based on the extent of the childhood trauma and the psychological and cognitive damage incurred. I discuss more about when it is appropriate to deploy the medical and moral modes in Chapter 6.

<sup>59</sup> While exemptions block any ascription of responsibility, excuses mitigate blame. More on the excuses and exemptions in Chapter 6.

challenge them as an adult. Even though mental illness and the trauma from one's upbringing are often intractable, so too are certain aspects of character. We should be careful not to overstate nor understate the intractability of these features of the self, as we should also be careful not to overstate nor understate the agency that adults (and some older children) have in being able to challenge the things they have been taught. Wolf's view understates the agent's influence and responsibility.

#### 1.4 Hardening Your Heart

Let us now turn to the third case, that of an agent who hardens her own heart. Assume that despite a good upbringing, in adulthood this agent slowly becomes entrenched in a corrupt and violent organization, justifying her participation to herself bit by bit while ignoring evidence that would show the organization to be immoral. After doing this for a sufficient amount of time, she becomes wholly loyal to this organization and refuses to consider any possibility that it is corrupt or that her actions are suspect in any way. In doing so, she has made herself willfully ignorant and intractable in her beliefs. And, as such, it seems that she has made herself less flexible in her thinking and less able to see the True and the Good (at least with respect to her loyalty to this organization). As such, on Wolf's view, she should be regarded as less responsible.

If, like me, you see similarities between this example and the participation of ordinary people in atrocious moral acts throughout history, you might worry that Wolf's account lets, for example, Nazis and slave owners off the hook. And this is precisely what Wolf takes herself to be committed to. Wolf (1987) claims that we should give less than full responsibility to Nazis, slave owners, and male chauvinists because their values "may have been inevitable, given the social circumstances in which they developed. If we think that the agents could not help but be mistaken about their values, we do not blame them for the actions those values inspired" (382). As Michelle Moody-Adams (1994) rightly argues, it is a strong claim to think that people in these historical contexts had no

other choice but to have the values they did. By the historical presence of those who resisted the Nazis and opposed slavery, we can see that it was not inevitable that one become a Nazi or a proponent of slavery. Look at how much attitudes have changed on trans rights and LGBTQ acceptance in our lifetime! The backlash is very real, but it is clear that culture is fluid, ideas can change, and we can actively contest the dominant cultural values.

We should also recognize the ability of normal, good people to engage in awful things. We can easily condone or justify to ourselves immoral acts and systems when they benefit us and when the pain and suffering of those exploited by the system are remote and hidden from view. For example, we often ignore the slave labor and unsafe working practices that produced a cheap garment when we are several steps removed from those practices and fail to look into that particular company's labor structures. As someone reads more and more propaganda and entrenches themselves in an ideology, it can become even easier for them to accept atrocious conclusions and commit morally blameworthy offenses. One does not have to be a psychopath to take a dark path.

It is a strange consequence of Wolf's view that an agent who makes herself less able to see the True and the Good can make herself less responsible. You should not be able to get off the hook for your actions by making yourself more evil. There is also some slipperiness in Wolf's definition of the ability to come to see the True and the Good. While the notion of ability is a complex one, it is difficult to ascertain whether someone who firmly disagrees with a true moral claim has the ability to see the True and the Good. If that person does come around to agreeing with that claim, then they clearly have the ability to see the True and the Good. But, if that person firmly disagrees until the day they die or only espouses the moral view when it is prudent, then it is unclear that they have the ability. This lack of clarity makes applying Wolf's view to particular cases incredibly difficult, as any time we point to someone who has mistaken values, we may never know whether they have the relevant ability that would make them morally responsible. If this is right,



then it is difficult to know when it is appropriate to blame someone, especially in cases where the person's actions are morally reprehensible. Perhaps we could blame those people retroactively when they change their views to morally correct ones, but that consequence of the view doesn't seem to fit well with cases in which we seem to rightfully feel resentment towards unrepentant offenders.

I suspect, contra Wolf's claim, that the ability to comprehend reactive moral responses is sufficient for responsibility in these cases—the ability to recognize resentment and gratitude, what the resentment and gratitude were responding to in the first instance, and what they call for in terms of action. This more readily allows for responsibility in the face of disagreement. Even the hard-hearted immoralist can presumably understand the moral reactions people have to her; she either ignores them or rejects them. We can also then unproblematically judge others to be blameworthy when they understand moral expectations and fail to revise their judgments in accordance with what is right. And this would not let entrenched Nazis and other oppressors off the hook.

As we have seen through these three cases, Wolf's condition that agents be able to form their values in accordance with the True and the Good is far too strong. It uses history as a way to obviate responsibility instead of to contextualize it.

## Part 2: Fischer and Ravizza's Historical View

Fischer and Ravizza's account of responsibility is in some ways related to Wolf's view, due to the emphasis on reasons-responsiveness, but it also draws on a Frankfurt-style process of taking responsibility. An agent is responsible for a reasons-responsive mechanism only if the agent has at some point in the past taken responsibility for it. For some of the same reasons that I disagree with Wolf, I also push back against Fischer and Ravizza's account.

## 2.1 Fischer and Ravizza's Historical Account

Fischer and Ravizza's (1998) "guidance control" view of moral responsibility conceives of moral responsibility in terms of responsibility for a mechanism that must be the agent's own and that also must be responsive to reasons (89).<sup>60</sup> The latter requires both "regular reasons-receptivity," or "a pattern of actual and hypothetical recognition of reasons (some of which are moral reasons) that is ... minimally grounded in reality" (90), and "weak reasons-reactivity," or the ability to translate reasons (though not specifically moral reasons) into choices, where, for any given choice, there is at least one possible world in which the agent chooses otherwise in response to recognizing a reason that was not present in the actual decision-making scenario (holding fixed the mechanism) (69). On this view, "moral responsibility is an essentially *historical* notion," because certain historical conditions are required to determine whether, say, an action is the agent's own (170).

These historical conditions matter in two kinds of cases: 1) "The first sort of example is one in which an agent freely acts (and exercises guidance control) at one time, and thus brings it about that he is incapable of acting freely (and exercising guidance control) at a later time" (195). These are tracing cases (as I discussed in Chapter 2, Section 2.4), such as when a drunk driver is determined to be responsible for their actions taken while inebriated, because they imbibed alcohol while they still were able to exercise guidance control. 2) The second kind of case involves "a wide range of intuitively 'responsibility-undermining factors,'" including brainwashing, hypnosis, or direct stimulation of the brain by a meddling scientist (197). When the agent's mental states are "produced in these ways, the mechanism that issues in the relevant behavior is *not*, in an important intuitive sense, the *agent's own*" (197).

---

<sup>60</sup> Fischer and Ravizza state that their account is compatible with a view that grants responsibility for nonreflective, habit or instinct-based mechanisms (85-89).

To deal with the second kind of case, Fischer and Ravizza claim “that some sort of *process* of taking responsibility is necessary for moral responsibility [and for (say), an action to become the agent’s own]. And (of course) a process is essentially *historical*” (200). This helps to ground the agent’s responsibility for their reasons-responsive mechanism, as:

When an agent takes responsibility, then, he obviously is *not* accepting responsibility for only those actions *whatever their source*; rather, he is accepting responsibility for only those actions which flow from a certain source. This idea can be framed more precisely by saying that *an agent takes responsibility for acting from a particular kind of mechanism* (215).

For Fischer and Ravizza, taking responsibility is a process we generally undergo during our childhood moral education that involves “three major ingredients:”

- 1) First, an individual must see himself as the source of his behavior in the sense we have specified. That is, the individual must see himself as an agent; he must see that his choices and actions are efficacious in the world. The agent thus sees that his motivational states are the causal source – in certain characteristic ways – of upshots in the world (210-211).
- 2) Second, the individual must accept that he is a fair target of the reactive attitudes as a result of how he exercises this agency in certain contexts.... If, however, a person resolutely shows no moral response or *appreciation* of the moral force of the attitudes we take toward him, then eventually we must concede that he is not an appropriate partner in the conversation: he has not taken responsibility for himself. In this case, we stop resenting him as a person, and begin treating him as we would a distasteful object or a dangerous (or annoying) animal (211-213).
- 3) The third condition on taking responsibility requires that the individual’s view of himself specified in the first two conditions be based, in an appropriate way, on the evidence (213).

This process can occur in an implicit, nondeliberative way (214). Agents are also generally expected to take responsibility for their practical reason as well as nonreflective mechanisms, as children are also expected to take responsibility for “behavior they didn’t pause to think about (or deliberate about)” (216). Once an agent takes responsibility, “it is almost as if the agent has some sort of ‘standing policy’ with respect to that kind of mechanism. Thus, when the agent subsequently acts from a mechanism of that kind, that mechanism is *his own* insofar as he has already taken responsibility for acting from that kind of mechanism” (216).

Fischer and Ravizza's view is in some ways a synthesis of Wolf's and Frankfurt's views—there is a requirement that one be reasons-responsive in the right way and a historical condition that requires taking responsibility.<sup>61</sup> Harry Frankfurt's endorsement or identification (1971, 1975) view of responsibility rests on the idea that we are only persons if we consider what we want our desires to be and form second-order volitions, which are desires that some first-order desire be one's will.<sup>62</sup> In order to form second-order volitions, one must have rational capacities to become critically aware of one's own will (1971, 12). Second-order volitions, or the self-identification with some set of desires as one's will, are at the core of moral responsibility for Frankfurt.

According to Locke and Frankfurt,

to the extent that a person identifies himself with the springs of his actions, he takes responsibility for those actions and acquires moral responsibility for them; moreover, the questions of how the actions and his identifications with their springs are caused is irrelevant to the questions of whether he performs the actions freely or is morally responsible for performing them (1975, 121-122).

If someone refrains from identifying themselves with their desires, then he is a “passive bystander to his desires and what he does” and is not morally responsible (1975, 121).<sup>63</sup> Fischer and Ravizza note that, while Frankfurt “frequently states that history does not matter to ascriptions of moral responsibility, he nevertheless talks of ‘taking responsibility,’ which arguably is a historical notion” (185, *ftnt* 31). Fischer and Ravizza's historical emphasis on taking responsibility can thus be seen as a development of a Frankfurt-style view.

---

<sup>61</sup> Fischer and Ravizza also briefly point to the kinds of cases Wolf is concerned with when it comes to a poor upbringing: “So, for instance, it is often thought that an individual's background can have a crucial impact on his subsequent moral responsibility. If an individual has been subject to significant sorts of mental and physical abuse as a child or young adult, this may well imply that the agent is not subsequently morally responsible for at least some of his behavior” (187).

<sup>62</sup> For developments of and alterations to Frankfurt's views as discussed in this chapter, see Frankfurt (1992).

<sup>63</sup> Schroeder and Arpaly (1999) contend that anti-identification is not the same as externality and charge that Frankfurt mischaracterizes the pre-theoretical phenomena he builds upon. I will leave aside questions about alienation, identification, and externality for now.

## 2.2 Taking Responsibility

Fischer and Ravizza seem to be primarily concerned with history as a condition that determines the presence or absence of responsibility. While I will not revisit tracing arguments here, I suggest that the historical requirement that an agent has taken responsibility for their mechanism is not, in fact, necessary for holding responsible in a variety of cases. While an understanding of history may be relevant in some cases for determining when we can appropriately apply the mode of character, I primarily think about history in terms of how it shapes some narratives of character. Additionally, I take it that our theories about moral responsibility should track our everyday moral practices—if they appear to be too revisionary, that is a mark against them. This generally rules out theories that depend on inaccessible metaphysical facts. Turn now to a few cases that suggest Fischer and Ravizza’s historical view may require revision.

First, imagine that I’m sitting and drinking coffee at my local coffee shop/workspace and a fellow patron decides to sit down and start hitting on me despite my protestations that I’m working and don’t want to be bothered. If I am trying to figure out whether this person is responsible, it does not seem likely that I should ask myself “did this man take responsibility at some point in his upbringing such that he is responsible for the mechanism that has causal efficacy on his current actions”? Requiring this sort of historically-distant mindreading seems too stringent a condition on responsibility.<sup>64</sup>

Second, imagine that a young woman considers herself to be totally passive in her life, seeing everything as determined by the positions of the stars. Instead of realizing her own complicity in

---

<sup>64</sup> Even when we intentionally try to remove anger, we tend to refer to other features of someone’s past or experience in order to reconceptualize. For example, as Hochschild (2012) discusses in *The Managed Heart*, flight attendants would use a variety of strategies to suppress their moral reactive attitudes towards disruptive passengers: “the passenger demanding, constant attention could be conceived as a ‘victim of fear of flying.’ A drunk could be re-conceived as ‘just like a child’ (25)”. These strategies are also often gendered, as women and customer service workers are expected to suppress their own discomfort to appease the whims of those around them.

pushing away her former best friend, she instead blames it on the movements of celestial bodies.

Fischer and Ravizza consider such a case and defend their view as follows:

The basic idea is that an individual who really does not see himself as an agent and a fair target for the reactive attitudes cannot be deemed genuinely active and morally responsible. In *not seeing himself* in a certain way, he *fails to be* a morally responsible agent. Lacking the required view of himself, he *is* essentially passive, buffeted by forces that assail him.... But an individual who fails to take responsibility (in our sense) is a bit like a sailor who does not believe his rudder is working; he allows the boat to be buffeted by the strong winds. He does not guide the boat. The boat's movements are no reflection on him; rather, they are entirely attributable to the winds (221).

Perhaps the boat's movements are no reflection on him, but the fact that he is holding the wheel and failing to recognize that he *is* guiding the boat may still reflect on him. If the young woman complains that her best friend has left her due to fate, it may still be reasonable to protest "but *you* were the one who said mean things to her—that's why she's no longer friends with you! Why can't you just accept the consequences of *your* actions?" The failure to take responsibility itself may sometimes be culpable.

Third, imagine a case in which someone only ever accepts the reactive attitudes of love and gratitude but never sees himself as an appropriate target for resentment or indignation.<sup>65</sup> Someone like Donald Trump, for example, or a chauvinist boss who blames his female secretaries for any mistakes he makes while taking all the credit for their successes.<sup>66</sup> This partial acceptance of the reactive attitudes may reasonably cut these people off from mutual and healthy human relationships,

---

<sup>65</sup> You could also have an opposite kind of case in which the person only accepts resentment and blame but never accepts love or gratitude. I do not think it would be reasonable to claim that the shift when the agent learns to truly accept love and forgiveness affects whether they count as fully responsible in the sense Fischer and Ravizza are concerned with.

<sup>66</sup> Gendered expectations around responsibility often place blame on women while exonerating and excusing men. Kate Manne (2017) connects this phenomenon to the idea of *himpathy*, or "the excessive sympathy sometimes shown toward male perpetrators of sexual violence" (197). Women are expected to take responsibility for men's sexual activities, while men are seen as somehow still innocent. As I was told in middle school at my fundamentalist Christian church, it was my responsibility as a woman not to ever be alone in a room with another man, because he wouldn't be able to help himself. At my private Christian school, strict dress codes were enforced to keep girls modest. One high school teacher even instructed all the girls to raise their arms one day, and anyone that showed any midriff was written up and punished. The express purpose of this policy was to protect the poor high school boys who couldn't control their lust. The message was clear: if I was raped because I had allowed myself to be alone in a room with another man or I had worn something revealing, it would be *my* fault.

but it does not necessarily seem to block the appropriateness of resentment. While the objective stance may be the most prudent for dealing with these individuals on a day-to-day basis, it doesn't seem inappropriate for someone to adopt the participant stance and call their behavior out—"you're a total asshole!" As I noted when discussing Wolf's view, you shouldn't be able to escape responsibility by making yourself a worse person or refusing to take responsibility.

Contra Fischer and Ravizza, I am much more concerned with how an agent's history connotes vital information about how the agent's patterns of choice and attitudes unfold within an ongoing, developing historical context. This historical information allows us to make rich interpretations about an agent's values, projects, and trajectory. Additionally, whether the agent has taken responsibility can also matter for how we assess their character. Things the agent has endorsed or accepted as belonging to them often lend to different interpretations than things that the agent has not endorsed or do not see as belonging to them (even if, at the end of the day, they do). History matters for coloring many stories we tell about character; our histories are not necessarily a precondition for being able to deploy the mode of character.

### 2.3 Lessons for a Positive View

The preceding discussion suggests a number of lessons for how history matters to narratives of character. First, history and context can help us understand how well you reasoned, given the information available to you, as well as how well you chose, given the options available to you. This is particularly relevant for cases of negligence and cases of coercion. However, history and context do not always exculpate. Sometimes you should have known better; sometimes you clearly had other options available to you. We may react markedly differently to the sheltered freshman who is just now encountering a world outside of his bubble than to the callous senior who has simply dug into all her bad beliefs instead of questioning them.

Second, history and context also matter for picking out patterns and narrative arcs. Are you improving yourself and inching closer to happiness? Are your current actions informed by ongoing, intergenerational trauma that especially colors your interactions with your parents and grandparents? Are you becoming a worse person despite your efforts to improve yourself? All of these questions are expertly drawn out by the recent movie *Everything, Everywhere All at Once* (2022), in which the main protagonist, Evelyn, must right the wrongs she inherited from her father and reconnect with her daughter, who has become an interdimensional villain in response to Evelyn's slights. The narrative arc that Evelyn takes is one of self-redemption through working to break negative patterns and fighting to create better family dynamics.

Third, our histories frequently point towards solutions to present problems. In dealing with trauma, for example, past events must be reconciled with so that patients can live in the present. A good understanding of past patterns and events can help us figure out exactly what is going on with our reasons and choices in a given situation and where they're coming from. It can also help us understand which situations might exacerbate certain kinds of behavior. All of these observations may reasonably inform our expectations of each other: "Don't spend so much time with your old high school friends, they make you a more callous person;" "I see now why you're reacting so badly to any criticism, but you've got to find a way to change it. Please go to therapy to deal with your inability to accept responsibility."

## Conclusion

In sum, while I claim that history is an important element for some narratives of character, I do not think of history as performing a primarily responsibility conferring or removing condition. Instead, history is often relevant for providing rich context to evaluate the patterns in an agent's reasons and choices over time. The story changes as the details change, and subtle shifts in the



details of a story can dramatically shift the interpretation of it. You can be the victim and the villain at the same time. You can be a prick who never takes responsibility and is all the more blameworthy for it. You can learn from trauma and grow by processing your past experiences.

In the next chapter, I present a view of responsibility for character in terms of narratives told within specific relational contexts.

## Chapter 5: Responsibility for Character

Why should we think about character as a kind of narrative? What is the range of narratives of character that we can tell? Why do we care about narratives of character?

In this chapter, I argue that character is a kind of narrative that we tell which is constructed from a number of different elements I have already argued should matter in a good account of moral responsibility: judgments, choices, and history. In Chapter 1, I set the stage for this project by placing it within a specifically relational, reactive context, in which the reactive attitudes such as resentment and gratitude represent the paradigmatic cases of holding responsible. The narratives of character that matter for moral responsibility, then, are those that embed and evoke these paradigmatic reactive attitudes (and other reactive attitudes around and between them) in their evaluative and emotional content. In Chapter 2, I argued that we are responsible for many things that are not under our direct control—choice alone is not all that matters for moral responsibility. In Chapter 3, I demonstrated that choice still matters in a way distinct from our non-voluntary attitudes and judgments. I suggested that both choice and judgment are necessary for a good account of moral responsibility. In Chapter 4, I claimed that history mattered for responsibility in many cases because it often shapes and sometimes changes our expectations and evaluations of others as responsible individuals.

Now, I would like to put forth a broadly virtue theoretic account of character that takes our judgments of character to be narratives (of a sort) of different sizes and scopes, which unify various details of choices, judgments, and history together with the evaluational and emotional import of the reactive attitudes, which serve as something like an interpretation. Narratives of character are in many ways like traditional narratives, but not in every respect. Like traditional narratives, they are told in particular relational contexts for particular reasons. These narratives of character are not told from a god's-eye perspective, nor do they have to be. They are always incomplete from that

totalizing, objective standpoint, but they can be complete for a given purpose within a specific relational context. Narratives are something that we actively tell for our own needs to mark which features are salient to our evaluations. At the same time, the narratives it is fitting for us to tell are also bound by other ethical norms, including accuracy and charity.

Character is thus necessarily unstable and incomplete—it is not a set of stable dispositions or some pre-existing essence that we are picking up on. At the same time, narratives of character are particularly helpful for picking out important truths about patterns of thought and action, and they serve as a useful guide for us in our relationships with others. They help us understand who is responsible for what and how we should respond to ourselves and each other. We often tell and re-tell these narratives of character to better grasp their import and process our own emotions and evaluations, which may develop over time.

In Part 1, I explicate the view. In Part 2, I consider a number of interesting upshots.

## Part 1: Narratives of Character

In this section, I lay out which aspects of narrative are relevant for narratives of character, provide examples of common narratives of character to illustrate the wide range of narratives we can construct, discuss the details that should be included in these narratives, and explain what it means to get the interpretation of those details right.

### 1.1 A Few Preliminaries

As I am arguing that character is a narrative (or at the very least narrative-like), I need to say something about what a narrative is. While narratives of character are likely distinct from traditional

narratives, they share a number of core features.<sup>67</sup> Peter Goldie’s (2012) account most clearly lays out the conception of narrative that informs how I understand narratives of character. Due to the similarities between Goldie’s account and my own, I have decided to use the language of narrative, though “narrative-like” might be more appropriate as a way of carving off other notions about narratives that do not apply in the case of character.

Per Goldie’s definition of narrative:

A narrative or story is something that can be told or narrated, or just thought through in narrative thinking. It is more than just a bare annal or chronicle or list of a sequence of events, but a representation of those events which is shaped, organized, and colored, presenting those events, and the people involved in them, from a certain perspective or perspectives, and thereby giving narrative structure—coherence, meaningfulness, and evaluative and emotional import—to what is related (2).

For narratives of character, we take a sequence of attitudes and choices, understood within context, and present those events with an evaluative and emotional import that evokes the reactive attitudes, most notably gratitude and resentment, but also the rich range of attitudes around and between them.<sup>68</sup> This process is active and constructive, done in response to matters of interpersonal import.

As Goldie notes,

thinking through a narrative, and narrating a narrative publicly, are kinds of *action*, done for reasons, and an account of these reasons can explain why someone thought through or related *this* particular narrative at *this* particular time in *this* particular way (perhaps distortingly, perhaps passionately but without distortion) (150).

As narratives are constructed and told actively and for specific reasons, the narratives of character that we tell can themselves reveal yet other features of our character by prompting us to develop a meta-narrative: “she’s so critical: she always mentions people’s flaws when she introduces them.”

---

<sup>67</sup> It’s unclear that there is a singular notion of a narrative in the literary sense. Schechtman (1996), for instance, suggests that conventional and traditional narratives are linear in form; however, there may be a cluster of overlapping forms of a “standard narrative” that are more cyclical or perhaps told out of order (103-104). I once had a therapist suggest to me that I think of progress on trauma-related issues as a kind of spiral moving upward, with each turn bringing me back to the same issues as before but seen in a new light and with new skills. What felt like regression was reframed as progress through an inventive narrative frame.

<sup>68</sup> These attitudes very plausibly include shame. Michelle Mason (2010) argues that shame is warranted when “one has violated some legitimate ideal of character” (418). For a reply to Mason, see Baron (2018).

These narratives, told at whatever level is appropriate for the present purposes, do not have to be intentionally and consciously constructed step-by-step nor publicly narrated. When developing narratives, “it is not as though there is first, a completed narrative, and then, second, an evaluation and emotional response to the narrative; rather, the evaluation and emotional response themselves infuse the narrative, shaping and colouring it” (11).

As Bernard Williams (1985) instructs us, narratives commonly employ thick evaluative concepts such as treachery, brutality, and courage that already run together fact and value (129). These thick concepts are applied based on “what the world is like (for instance, ... how someone has behaved)” yet “their application usually involves a certain valuation of the situation, of persons or actions” (129). Moreover, they tend to “provide reasons for action” (129-130). Narratives of character, too, give us reasons to change how we view, react, and interact with other people—perhaps a friend’s dishonesty sours the friendship or a partner’s loving patience allows the relationship to deepen.

Even though narratives are perspectival and inextricably bound up in practical reasons, narratives can still be accurate or inaccurate, charitable or uncharitable.<sup>69</sup> Goldie claims that:

A narrative can be objective in this sense: it can be *appropriate*, involving an appropriate evaluation of, and emotional response to, what is related. So it is possible for a narrative to be objective, albeit emotionally engaged and perspectival (155).

I will say more about this in sections 1.3 and 1.4. There, I will discuss the appropriateness of narratives of character through my own language of “details” and “interpretation” to distinguish between the sequence of attitudes and actions understood in context and the evaluative and emotional import we assign to that sequence, respectively. In order for narratives of character to be appropriate, they need to select the right details, not omitting relevant details nor including irrelevant

---

<sup>69</sup> When we tell narratives about ourselves, Goldie thinks that we take an external perspective on ourselves, “and, in just this respect, one is at the same time both actor and spectator, both narrator and audience, and both agent and judge, judging both the events in which one’s past or future self is implicated, and judging the narrative itself” (42).

ones (given the purpose at hand, with some flexibility).<sup>70</sup> They also need to get the right interpretation of those details (or, at least, one of the right interpretations in a range of acceptable ones). The details help us understand the particulars of how someone is racist, kind, transphobic, or inclusive. The interpretation gives us the evaluative and emotional language to make sense of how the particulars are unified. As Gregory Currie (2010) notes, the selections we make in our narratives of character indicate that the information the narrative imparts “is Character-relevant to our concerns—worth the processing cost of calculating what it implies about Character” (190-191).

Finally, narratives of character do not need to reveal a consistent character trait or disposition.<sup>71</sup> Sometimes they reveal deep inconsistencies or sets of conflicting attitudes that sit at the edge of what we can psychologically understand as fitting together in the same person. And not all actions and attitudes need to be fully explainable by the agent in a way that coheres—sometimes people act akratically, against their best judgment, and they can’t give any good reason why. Perhaps they simply acted differently on a whim, or it just happened that way somehow. These features do not necessarily exclude these events from narratives of character.

## 1.2 Common Narratives

Narratives of character come in a wide variety of forms. They can span a single incident or a full lifetime; they might be reasonably told out of order or may need to be told in the exact order that everything happened. In this section, I’d like to draw our attention to a range of common narratives of character that differ in their form and scope.

---

<sup>70</sup> I take it that details can include what Goldie calls “general events,” which are reoccurring events such as a parent taking a child to a playground or two people getting into the same argument over and over (46).

<sup>71</sup> This firmly places my account outside of the view that our moral character is an element of narrative, understood as a kind of central, stable, and motivating set of dispositions. This notion of literary character may be a useful simplification for plot development, but actual people are rarely as straightforward as fictional characters. I think it is possible to tell narratives of character of the sense I have in mind about fictional characters, but those look much more like the kinds of nuanced analyses that bring together a number of facts about the character’s background, attitudes, and choices to construct an interpretation that may be argued over.

As a first example, take a single instance in which a rude customer spouts off at a service worker for no good reason. Say, for instance, that a customer repeatedly insists that she should be able to have a large party of people seated immediately on a busy night at a restaurant without a prior reservation and only escalates her behavior when she is denied. The service worker who interacts with this customer can potentially construct a narrow narrative of character based off of this single interaction, should she have sufficient reason to do so. (As I'll discuss in section 2.1, this is a borderline case of a narrative of character due to its very limited scope, but I'll grant for now that it's plausible as a narrative.) This kind of narrative is a circumscribed account of character that includes limited information, but the contextual details may still be enough that it's reasonable to assume that the customer is entitled and to respond accordingly. At the same time, the limited nature of these snapshot narratives can mislead us in our interpretations. Perhaps someone is having a particularly bad day or at the start of a manic episode.<sup>72</sup> Those details might change the narrative dramatically. However, epistemic humility does not require total skepticism about whether we can tell accurate narratives in any instance. Sometimes the details are clear enough, even at first glance.

As a second example, take gratitude within a close friendship. One of my very close friends has repeatedly helped me reinterpret how I see myself in a more loving light. I am particularly bad at accurately and charitably assessing my own motives and actions, as I tend to over-emphasize my faults and under-emphasize my virtues. I will always be grateful for that friend's insights, which helped me cut through the ever-present self-critical voice and dampen its influence. This is a wider narrative of character than the first, in that it spans a larger pattern unfolding over time. It may intersect with other narratives I could tell about this friend's character, but it can also stand on its own, depending on the context and which stories need to be told. Note, however, that in other

---

<sup>72</sup> The stance that requires us to suspend all judgments because of these potential excuses can be quite demanding and place an unequal burden on us for managing these kinds of interactions and our own internal reactions when the other person has not managed themselves well.

contexts I may tell other stories about this friend's character for other reasons. These narratives are not competing accounts; they are complimentary accounts constructed for different purposes.

As a third example, take the cancelling of J.K. Rowling. If you're explaining to someone why J.K. Rowling is transphobic, there are certain instances that need to be relayed, including her personal testimony, her tweets, her recent novels, who she defends and associates with, etc.<sup>73</sup> Most retellings of the story about J.K. Rowling include roughly the same core details, understood in sequential order and updated with each new incident. Rarely does anyone feel the need to dig up every detail of Rowling's life to try to make a "complete" assessment. Even ContraPoints, who made an [extensive video on the subject](#), focused on the details of Rowling's history that might bear on her current transphobic stance. The less in-depth assessments are still enough to tell us not only that Rowling *is* transphobic but also *how* she is transphobic.<sup>74</sup> On the basis of these narratives, we can hold Rowling accountable for her transphobic judgments and choices. Note that resentment is a relevant attitude here, but so is feeling betrayed or disappointed, especially for those in the trans community who grew up with Harry Potter.

As a fourth example, take the podcast *You're Wrong About*, which tells extensive accounts of character through deep dives into the stories of people's lives. This podcast re-tells a series of stories about the lives of [maligned women in the 90s](#), engaging in sensitive biographical work aimed to more accurately and charitably interpret these women's lives. These rich stories are almost always mixed and complex, and they tend to elicit a wide variety of evaluative and emotional reactions in us. These stories also tend to have the richest historical context and background, which is decidedly

---

<sup>73</sup> One such article that lays out a fairly comprehensive history can be found [here](#), though not every telling of the story needs to include all of those details.

<sup>74</sup> Cancellations are also a place where the interpretations of the details of someone's actions can go wrong—sometimes too much is inferred from a single instance that contravenes the rest of the pattern. See Sarah Schulman's (2016) book *Conflict is Not Abuse* as an argument about trauma leading to overreactions in a similar way as supremacy mindsets, especially when it comes to paranoid assumptions that people are either good or bad and that the bad people tend to hide behind masks that can be seen through because of one single bad action.



relevant for our reactions and assessments. For example, a friend of mine started reading a biography of Fred Rogers, who he looked up to, only to discover that Rogers was born into money and never suffered hardship. This, I think reasonably, changed my friend's feelings toward and assessment of Rogers' goodness and accomplishments. We frequently have these very complicated relationships with other people, deeply admiring them for some features, resenting them for others, and sometimes finding ourselves at a loss for how to respond. Narratives of character allow us to hold these tensions together and try to respond to the whole that emerges.<sup>75</sup>

The preceding four cases are all stories about character told for different reasons at different times. They stand on their own grounds with their own sets of details—they do not need to be derived from some complete story of character that encompasses every potentially morally-relevant detail in a person's life. They have not only their own sets of *details* but also their own *interpretations* that fall out of the details. For example, J.K. Rowling's actions suggest an interpretation of transphobia. But note that both the details and the interpretation are needed to make sense of character. If we lose the details, we lose the particular ways in which Rowling is transphobic. If we lose the interpretation, then we don't see how the details line up to reveal transphobia.

### 1.3 The Details We Care About

What details matter for narratives of character? What elements can be included in stories about character and moral responsibility? Something like the following list:

- our judgments/attitudes/reasons/cares/values,<sup>76</sup>

---

<sup>75</sup> George Eliot's *Theophrastus Such* (2023, originally published in 1879) contains one such delightful tale of “a man declared to have a bad temper and yet glorified as the possessor of every high quality” (34). The narrator dutifully notes all the myriad contradictions in the man as well as how frustrating it is to deal with his capriciousness and hollow excuses.

<sup>76</sup> I take it that this includes our emotions insofar as they are judgment-sensitive and responsive to certain features of the world. For views on how we should understand the emotions and their evaluative content, see Deonna (2006), Betzler (2009), and Prinz (2004).

- our choices/things we reflectively endorse,
- our histories and present context, and
- other aspects of what we're like, including our social identities, personal conceptions, embodiment, etc.

This list is meant to be pluralistic. Each element in this list might be broken down into finer-grained distinctions, but we do not tend to distinguish between, say, values and reasons when telling our everyday stories of character.<sup>77</sup> When we deploy the mode of character, we are trying to say something relatively holistic about real people and their complexities.

Though some element from the first or second bullet point is required and sometimes sufficient to get a narrative of character off the ground, narratives of character tend to presuppose the last two bullet points, which can dramatically change how we understand the reasons an agent recognizes or the choices they make.<sup>78</sup> With regards to choices and judgments, I do not think there is a clear priority between the two—things under our direct control are not necessarily better nor more truly ours than those under our indirect control. Our choices and our judgments simply matter to us in different ways.

As I argued in Chapters 2 and 3, both our choices and our judgments (which may be non-voluntary) are necessary to capture the unfolding patterns of activity and passivity within our relationships that are commonly targeted by reactive attitudes such as resentment and gratitude. The

---

<sup>77</sup> My account is silent on the question as to whether reasons themselves are motivating. The level of description that matters for our moral assessments of ourselves and others is much coarser grained than the kind of description that is necessary to explain how people are motivated by the content of their reasons. See Korsgaard (1996b) for an account that takes reasons themselves to be motivating. See Arpaly and Schroeder (2013) for an account that gives an important role to desires for moral motivation.

<sup>78</sup> I also don't think the solution here is to try to fit the last two bullets within the definition of the first two. I am in general skeptical of reductionist models. I'm specifically worried that by building in our histories, context, and identities into our reasons or into our choices that still we wind up constructing an isolated, rational agent and obscure how embedded we are in the world around us. There are also all sorts of complex philosophical questions about the relationships between the different elements represented in the bullets—are our reasons and choices themselves a feature of embodiment, or are they tied to a mind that is separate from the body? These questions would take us too far afield for the present purposes.

complex interplay between what is under our indirect control and what is under our direct control can produce a number of different narrative structures that bear on how we respond to agents engaged in internal strife: Does the person find themselves with entrenched bad attitudes that they are actively working to remedy? Did they intuitively know the right thing to do and choose to do the wrong thing anyway? Are they inconsistent in their choices, only sometimes working to become a better person? Are they internally conflicted at the level of their attitudes but don't realize it?

The third bullet point, our histories and present context, is often needed to make sense of how a pattern developed, or what information someone had access to and which actions they could have reasonably taken in a given situation. As Bernard Williams (1981) states, "one's history as an agent is a web in which anything that is the product of the will is surrounded and held up and partly formed by things that are not" (29). Our histories and context help us answer questions like: Is a daughter's intense reactions to her mother inappropriate given their long history, in which the mother has said many nasty things to the daughter and still fails to respect her as much as she does her other children? Is resentment a justified response to a roommate's failure to adequately contribute to cleaning, or did a long history of enmeshment in the resentful roommate's family prime them to become codependent in the relationship?

The moral communities we find ourselves to be a part of help to constitute our context and history. The power structures and institutions we live within, our support systems, and the people we care about can create real limits and hurdles to us when trying to live out our values or do what is good and right. Just as we shape our moral communities and the cultures around us, so they shape us. Though I've placed our social identities and personal conceptions in the last bullet point concerning what we are like, both are strongly shaped by our communities. In the case of social identity, one sense of it is almost exclusively shaped by community: the sense of how we are perceived within a given context. Both this sense and the sense of how we conceive of ourselves can

be relevant for narratives of character. For example, it's very different to tell an edgy non-binary joke as a non-binary person than as a cisgender person. But sometimes social identity makes our assessments of character more biting, such as our assessments of log cabin Republicans.

When it comes to our broader self-conceptions, they are as much a part of the details others might employ in a narrative about us as they are themselves a kind of interpretation that we construct for ourselves. If someone sees their flaws accurately yet lacks any desire to remedy them, we tend to respond differently to that person than to someone who hasn't yet seen the problem. The narrative and its evaluative import is relevantly changed. Note that we might have multiple, competing self-conceptions and that if something is contained in one's self-conception, it is not necessarily endorsed.

Finally, embodiedness appears to inform all the prior elements, at least when we're talking about the character of human beings.<sup>79</sup> We often want to place this piece away from our stories about responsibility, treating conditions like being hormonal or hangry as somehow wholly separate from our responsible selves. It is unclear to me that these conditions can be so easily separated, as they frequently modify our assessments of how deeply we hold certain attitudes or why we made certain choices. Tiredness, drunkenness, and rich experiences of disability can matter for our stories of character, and not necessarily in a way that removes responsibility. See, for instance, the assessment that "he's a mean drunk." That appears to be a character judgment, despite the fact that the man's drunkenness would tend to remove responsibility on many other contemporary accounts. The meanness is likely something that is already there that the man represses, but it clearly rears its

---

<sup>79</sup> Greco (2015) notes that Hume also recognizes a central role for the body: "If one then considers the catalogue of mental qualities constituting virtues and vices given in Book 3 of the *Treatise* and in the *Enquiry concerning the Principles of Morals* (see esp. T 3.3.4–5; EPM 6–8), one clearly sees that they can be framed only by conceiving human beings in concrete circumstances in which they are engaged as corporeal agents. As a matter of fact, Hume already recognizes a central role for the body in the definition of the person in Book 1 (see T 1.4.2.9/ SBN 190; and T 1.4.5.30/SBN 248). And, more generally, it is the very idea of human beings as feeling creatures that cannot hold if they come to be understood solely as mental beings" (709).

head when he drinks.<sup>80</sup> Or, take an older woman who undergoes anesthesia and, while waking up, says all sorts of mean things to one of her children while saying nothing of the sort to her other two children in the room. Assume that this very closely aligns to other patterns of behavior that she has exhibited while perfectly sober and alert. This may be an especially clear moment in which her views are laid bare, even though she normally tries to keep them to herself. The drunkenness or anesthesia changes the narrative and our assessment of the agent's attitudes and choices, but it does not seem to preclude telling such a narrative about the mean drunk or feeling resentment in response to the mother's comments. In this way, embodiment is itself part of the history and context that helps us to interpret the attitudes and choices expressed.

As noted, none of these categories can be neatly cleaved from each other, and many actively shape each other. Social identity makes up part of your present context, and the present context is already informed by history. Choices are made within a set of available reasons within a context, and all of this is mediated by embodiedness. Interestingly, not all elements are necessary for each narrative of character—depending on practical needs, someone's history might either be irrelevant or take center stage. The stories we tell about character are constrained by relational needs and by our practical judgment. In the process of crafting narratives of character, we make judgments about what is character-relevant and necessary for understanding the pattern or incident in question. For this reason, choosing the details to include is itself already tied up in interpretation.

---

<sup>80</sup> These cases can be exceptionally complicated due to the varied ways that people use and are affected by alcohol and other substances. The point is that if we carve off a whole set of bodily states, including drunkenness, as necessarily responsibility-exempting because the medical mode is appropriate, then we may very well miss important patterns of evaluation and choice that are implicated by the mode of character.

## 1.4 Getting the Interpretation(s) Right

Our moral interpretive language is rich and varied, including the traditional language of virtues and vices, as well as a vast range of other interpersonal patterns that we have named, such as enmeshment, gaslighting, active listening, nit-picking, etc.<sup>81</sup> Our interpretations of the patterns in others' choices and attitudes (understood within context) make use of this general language, but the details of the story help us understand *how* someone is cowardly or courageous, callous or kind. A pattern of codependency, for instance, may look markedly different from one relationship to the next.

Our interpretations can also involve multiple emotional and evaluative valences and be quite complex—they are often interestingly mixed.<sup>82</sup> Narratives are rich structures that are perhaps more like poetry than cold, analytic philosophy; they allow for multiplicities of meaning and new interpretations. Unlike some poetry, however, we can directly and clearly communicate the core evaluational and emotional content that a narrative evokes. In constructing and interpreting a narrative, we may never say everything that there is to possibly be said about it. But we can capture the core details and importance of the story into something that we can tell and re-tell, feel and feel through again and again. For this reason, narratives are not stable and complete, nor should they be. As our relationships develop and we glean new information, the narratives we tell about character should shift and change.

---

<sup>81</sup> I also leave open the possibility of burdened virtues that are good in particular contexts but detrimental overall to the agent's flourishing. As Lisa Tessman (2005) argues, some kinds of anger can be necessary for social change but also corrosive to the bearer. Myisha Cherry (2021), by contrast, provides the example of Lordean anger as a kind of anger that avoids these problems.

<sup>82</sup> As Pamela Hieronymi (2019) notes, because the reactive attitudes “are not sensitive to every aspect of a person's situation, a person cannot be justified in them by appeal to just any aspect of the situation, nor, importantly, can their justification be challenged by just any aspect. These attitudes embody your take, not on the all-things-considered question of whether it would be best to have those attitudes, but instead on some narrower set of questions which concern (not the attitude itself, but rather) only certain aspects of your situation. And so the question of whether you have reason to resent, e.g., and therefore the question of whether you are justified in resenting, cannot be answered by appeal to just any aspect of your situation or relationship. It can be answered only by appeal to a narrower, more specific set of considerations” (13). For this reason, it makes sense that we would feel more than one reactive attitude in complex cases.

Interpretations are good when they are accurate and charitable. But part of what makes an interpretation good or not is whether the right details have been included. Which details count as the right details is further determined by the practical reasons (which can be legitimate or illegitimate) behind telling *this* narrative of character. The particular epistemic requirements on a narrative of character in any given situation depend strongly on the relationship, the context, the significance of the action or pattern involved, and a host of other considerations. Some general norms, such as respect for others and fairness, dictate at the broadest level how we should select the relevant details. But neither fairness nor respect for others require a god's-eye view or perfection in storytelling. Brief, one-off encounters may be sufficient for very limited narratives that capture that particular instance apart from the broader pattern. More in-depth relationships tend to warrant much deeper and more extensive narratives of character.

The specific reasons within the situation, such as an immediate need to name and understand the problem that's causing friction in a romantic relationship, draw our attention to specific details like the fights that have occurred or the moment at which things appeared to turn sour. But these details are not fully pre-set. We may try to decide whether a particular historical fact about a partner's childhood has any bearing on the present situation or if the irritation the partner exhibited two weeks ago was due to the ongoing conflict or something unrelated. If I am right that we are the ones who actively draw these features together into narratives, there is not some pre-determined narrative or narrative structure that we are picking up on—we are creating it ourselves.

At the same time, we can absolutely tell false narratives of character and get the interpretation wrong. In cases in which we have very limited information about the details involved, we may have a difficult time getting the right interpretation: Is this person just missing social cues? Are they maybe also an asshole? Is someone grieving a recent breakup and stuck in their own world?

Or do they genuinely not care about other people as much? Is this person acting for the best interests of others like they say they are, or are they secretly motivated by their own selfish ends?

We also have tendencies to over-vilify and over-valorize individuals. If we oversimplify character into some reductive, stable disposition, we can easily flatten out the characters of others and treat them as if they were fictional villains or heroes in a way that casts all our interpretations of their character in a particular light, regardless of the details.<sup>83</sup> Goldie notes that “as a result, we tend to expect too much of our heroes and too little of our villains” (169). On the one hand is something like the “bitch eating crackers” phenomenon, in which everything the person does becomes reason to evaluate them negatively, regardless of whether that evaluation is appropriate, as when someone is blamed for eating crackers the wrong way because they are already so resoundingly disliked. On the other hand is the over-valorization of someone as a “good guy” who couldn’t possibly act from bad motives or choose wrongly. Of course, there are quite a number of inaccurate responses around and in between. Though literary fiction and non-fiction can oversimplify our expectations when it comes to character, behavior, and narrative interpretation, a diet of good, nuanced, and varied narratives can improve our ability to construct good narratives of character.<sup>84</sup> If we are not subtle enough with our own narrative vocabulary and ability to handle nuance, we can quickly fall into inappropriate evaluations and emotional responses.

There are, however, a few quick tests we can employ to try to avoid some of the common pitfalls: A) Is the interpretation backed up by specific events? If a co-worker is supposed to be a master manipulator, then what exactly has she done that suggests this interpretation? If no details can be cited in support of the interpretation, then the interpretation is likely wrong. B) Are the

---

<sup>83</sup> There can also be a cynical tendency in thinking that everyone is equally bad, or perhaps equally both good and bad.

<sup>84</sup> Good literary analysis can help us develop good narratives of character about fictional characters, by unifying the particulars of that character’s story into a narrative with a comprehensive evaluative and emotional import. A paucity of narrative tropes can also leave us with inadequate resources to understand our own lives—for example, the simplistic testimonial narratives of Christian conversion that I grew up with demanded a particular form, even if my life did not fit into it.



individual events in the story accurately represented? Or have we entered “bitch eating crackers” territory? C) Do the events and details, taken as a whole, align with the interpretation given? Is the “good guy” whose bad behavior is constantly being excused really a good guy? These tests are not exhaustive.

At this point, you may be wondering—where is any mention of habit? I suspect that the notion of habit cuts across both the details and the interpretation. If a particular action was done out of habit in the sense of being done on autopilot, then that may be a relevant detail. If a series of actions and reasons are being strung together into a recognizable pattern, then habit is already happening at the level of interpretation. Something like Julia Annas’ (2011) conception of virtue as similar to a skill, which is intelligently applied in action instead of merely routine, can also be captured at the level of interpretation. Importantly, however, I do not want to limit our narratives of character to only deal in habits—our non-habitual actions can say quite a bit about who we are and what we care about, and they frequently warrant the reactive attitudes.

Finally, on my definition of character, “acting out of character” has to do more with narrative coherence than the stability of a disposition. If someone is inherently erratic, it may be difficult to find an instance that breaks the pattern. However, if someone generally acts and reasons in line with a regular pattern but contravenes it only once or twice, it may be appropriate to say that they are acting out of character in the sense that the outliers don’t fit with the standing interpretation we had of their previous actions. Sometimes the actions that contravene the pattern suggest that the interpretation should be changed, but in other cases it may be possible to have a few incidents that don’t square with the overall interpretive frame yet do not displace it.

## Part 2: Interesting Upshots

In this section, I discuss a number of relevant upshots, including the idea that character is not necessarily stable and complete, that we can be responsible for some things that we appear to be passive towards, that narratives of character are not the same as our lives, and that narratives of character can still be factual and accurate.

### 2.1 Incomplete, Unstable Character

Character is often conceived of as a set of stable dispositions or, when understood as a narrative, something that necessarily extends throughout a person's entire life.<sup>85</sup> If my view is right, it is possible to tell narratives of character that reveal a set of stable dispositions or that span a person's life.<sup>86</sup> However, narratives of character have a much broader range than these narrow definitions allow—they can be incomplete (relative to the standard of a whole life) and reveal someone to be unstable and inconsistent. On my view, no person has a single narrative of character that attaches to them—different narratives may pick out different features for different reasons, and they may all be reasonably accurate and encompass the details needed to make sense of the particular pattern or instance of it. Nor are these narratives derived from one master narrative.<sup>87</sup>

---

<sup>85</sup> Alasdair MacIntyre (1981) is the prime example for the latter kind of view.

<sup>86</sup> As Rachana Kamtekar (2004) explains, thinking about character as excessively stable and situation independent gives ammunition to those who would reject the idea of character altogether: “the experiments which find character traits to correlate poorly with behavior rely on a very particular conception of a character trait: as an isolable and nonrational disposition to manifest a given stereotypical behavior that differs from the behavior of others and is fairly situation insensitive.... According to traditional virtue ethicists, virtues are not dispositions to perform stereotypical actions popularly associated with a given trait but, rather, dispositions to respond appropriately—in judgment, feeling, and action—to one's situation. Such responses require the active involvement of the agent's powers of reasoning” (477). Julia Annas (2005) concurs that the situationist misses the point that virtue is a like flexible and dynamic skill that becomes less habit-driven the more virtuous you become.

Similarly, Elizabeth Spelman (1978) worries that viewing character as situationally determined fails to treat other people as persons: “Just as we cannot be sure that we are treating persons as the persons they are by following moral rules, so we cannot be sure that we are treating them as the persons they are if we see them as beings whose character traits and behavior can be subsumed under psychological and sociological laws” (157).

<sup>87</sup> We could combine the different narratives into a broader narrative; however, there is not one single master narrative that grounds every localized narrative.

There are interesting questions about how bare bones a narrative of character can be and still count as a narrative. This is a tricky question that doesn't admit to some clear threshold or set of necessary and sufficient conditions. Remember that Goldie defines narrative in terms of a "representation of those events which is shaped, organized, and colored" (2). This suggests that narrative requires something more than a simple event—events *plural* are needed to construct a narrative of character. In the case of narratives of character, I will assume that, at the very least, more than one attitude or choice is necessary to construct a narrative. This leaves the example of the barista's interaction with the customer in question. If the interaction is sufficiently drawn out to involve more than some simple event, then there may be enough information to string together the attitudes the customer expresses and the choices they make to construct a very minimal account of character. If only one attitude is expressed or one choice is made, we may be forced to make a much more minimal judgment about that singular choice or attitude. That choice or attitude may still be colored with some evaluative or emotional import, but it lacks the narrative scope of character.

Additionally, since narratives of character are told in particular contexts for particular reasons, the stories we tell at different times can pick up on different features of our lives and emphasize different aspects, depending on the needs at hand and new events that might inform how we understand the old ones. As Lorenzo Greco (2015) interprets Hume's view of character, the "perspective on the self is always 'from here' ... this narrative is always open, as it were, always liable to be rewritten from the present moment in which we become self-conscious" (715-716).

Narratives of character thus lack narrative closure—they are always open-ended. Even if we are retelling narratives about the character of people who have died, we may find new features salient at different times and find our stories developing as we grieve. At some point we may decide that we have told and re-told a particular story long enough, but this closure is something we bestow. It is not strictly set out by narratives of character themselves, though the relative coherence

of a narrative may indicate whether we are ready to leave it behind. If it doesn't yet fully make sense, we may revisit it again and again until either it does or until we accept that it doesn't make sense. Sometimes it comes time to put old wrongs behind us. As Goldie notes, "it is no easy matter to live with someone who incessantly tortures himself with his past wrongdoing" (103).<sup>88</sup> Nor is it easy to live with someone who will not let go of any petty resentments or offer forgiveness.

## 2.2 Activity and Passivity in Moral Life

Goldie argues that, in our narratives of character, we are not commonly concerned with giving a "causal account or explanation of a particular event, but rather appeal to a kind of pattern of activity and passivity" (47). If my account is correct, then we can be responsible for things that we are passive with respect to. The things that occur to us, the reasons that come to mind, the features of situations we notice, and the reactions we find ourselves with are all (at least relatively) passive features of our moral lives.<sup>89</sup> And yet, as I have argued, we are responsible for them as part of our judgments/values/cares/reasons. At the same time, these states can't capture the importance of direct control and choice. Both matter in different ways for moral responsibility.

We are also inextricably perspectival and embodied, living within a particular historical context that shapes and delimits our ability to know and understand and choose. As members of the moral community, we make claims on each other, treat each other in ways that communicate respect

---

<sup>88</sup> Hieronymi (2019) suggests that in these cases, the person blaming themselves has missed the fact that while the criticism is *of* you, it is *about* the person who is raising the criticism: "In fact, you are missing the most important aspect of your situation: the fact that there is another person who has been disregarded, whose own importance has been overlooked, and who therefore now needs your attention, apology, and redoubled concern. By falling into disappointment and self-reproach over your own failure, you risk once again overlooking the importance of the other person" (25-26).

<sup>89</sup> Smith (2008) contends that these should be counted as active features of our moral lives so long as they fall within the scope of evaluative judgment and should be sensitive to our evaluative judgments and commitments (263). "Nothing much turns on our use of the terms 'active' or 'passive' here, however. For those who associate 'activity' closely with deliberation and choice, my view implies that we can be responsible for things with respect to which we are (volitionally) passive" (263).

or disrespect, and we are often partially at the mercy of others when it comes to the development of our values and cares, as well as the choices that are open to us. A particularly cutting comment from a close friend can stick with you for quite some time and alter how you move in the world. A manipulative parent can withhold key financial resources from an adult child to control their behavior. We can also be made to see our own worth through the kindness and respect of a community that elevates people like us when few others will. Close friendships tend to adjust how we see the world and which values are alive to us. None of this means that we are responsible for others, but it does mean that our actions can influence other people in ways that shape their character and that we can be likewise influenced, in part by the narratives we tell about ourselves and each other. This is the cost and promise of living within a moral community.

We could try to limit responsibility to the set of things we can directly control, but this so limits the things we can have reactive attitudes towards that it is radically revisionary of our ordinary moral practices. However, if we see the stakes of moral responsibility as primarily pertaining to our relationships with others as opposed to something like eternal damnation, there is often the possibility of repair, given actual contrition and forgiveness. If we are not legalists in our relationships, cutting people off at any sign of fault, we can learn to trust others to be kind and forgiving to us as well for our own foibles.

Thus, we have a picture on which moral persons are both passive and active in moral life. We often react and respond to the world and each other in ways that we can only indirectly control, and yet we also choose to act in ways that shape who we are. We are passive with respect to many of the ways that our situations and the communities around us shape us, but we can choose how to respond to the situations we find ourselves in and decide what we want our moral communities to look like. Both facets of moral life are necessary to make sense of character.

### 2.3 Character as Narrative; Not the Essence that Underlies the Narrative

Why say that character itself *is* a kind of narrative rather than that we tell narratives *about* character? As Goldie notes, “a narrative is distinct from what it is a narrative *of*” (6). In my view, character is a narrative representation that uses a rich vocabulary of vices, virtues, and other relational patterns to communicate things about a person. To say that character just is the set of actions and choices understood in context is “to run together what is represented with the representation” (153-154) and “to assimilate language and the world” (153).

On this view, character is not some deeper essence that we directly pick up on; it is an actively constructed representation of us as persons that is made for particular situations and practical reasons. These representations help us make sense of those situations and filter our information to make it easier for us to process salient aspects of someone’s psychology. We have to draw together the language and concepts to understand ourselves and others and the relationships we have. Character happens at the level of us already embedded as members of the moral community at the level of moral thought—it is not some metaphysically prior essence. So long as the narrative accurately tracks choices and attitudes in the world and is complete for the purposes at hand, we do not need to, as it were, peer under the metaphysical hood to see if there is an independent essence that the story picks up on. Essence is something much slipperier and deeper that we do not always have an excellent grasp on. We need enough of a grasp on the actual person’s reasons and choices and feelings to create an accurate narrative of character, but that’s a much less demanding task than determining their (even localized) essence. Character doesn’t have to be something real and hidden that our theorizing tries to magically reveal. When we tell different narratives and evaluate those narratives in a critical lens, this practice constructs character. If we

didn't tell these kinds of stories and reflect on them together, the notion of character would not have application.<sup>90</sup>

We are intentionally taking our very complex selves and representing them in ways that make ourselves more comprehensible to ourselves and each other. These representations need to be charitable and accurate and complete for the purpose at hand, and they can still do all the work we want in identifying patterns relevant to gratitude and resentment (and other emotions around them) without having to appeal to some mysterious deeper essence that is dubiously there in the first place. Through narratives of character, we give meaning to ourselves.

Making out character to be a narrative also provides us with a kind of agency that the stable disposition/essence view sometimes denies. We exercise agency in telling and shaping narratives of character about ourselves and others—we can practice finding the good in others and being encouraging while also constructively critical. We can also change our own actions and attitudes over time to shift our own narratives of character. Since our character is mutable, there is always the possibility for change, even though it may be exceptionally difficult in many cases.

As Goldie notes,

Of course, the simple fact that our lives, or episodes in our lives, are not narratives (that is, they are not *identical* to narratives) does not imply that narratives and narrative thinking cannot play a central part in how we lead our lives. We think, talk, and write *about* our lives by narrating or thinking through narratives, and how we do this can profoundly affect our lives as such. What emerges, then, is that our narratives of our lives, or of segments of our lives, can be embedded in the lives that we lead, which themselves are not narratives (6).

Character, even when understood as a kind of narrative, is still central to our practices of moral responsibility.<sup>91</sup>

---

<sup>90</sup> Thanks to Ricky Mouser for putting this point exceptionally clearly.

<sup>91</sup> Some of my language may be moderately revisionary when considering our ordinary language. We might say something like “I was ungenerous in my assessment of her character.” I think this is a loose way of speaking, but we can say something more precise like “I was ungenerous in my assessment of her” or “I was ungenerous in my interpretation of her actions” on my view. I don't think much hangs on whether every ordinary statement about character can be accommodated on my view. At a deeper level, you can absolutely be ungenerous and mistaken in the narratives you tell

## 2.4 Relativism? and Narratives of Character

How relativistic is this view of character? In Goldie's words, "factual narratives can be true *period*, and there is nothing in the nature of narrative that excludes this possibility" (150). Our narratives of character are essentially perspectival and bound up in the view from where we are at the moment, but they can still be true. We have to be in the middle of the activity of telling stories about character to make sense of it at all. Even if something like a god's-eye view of character exists, it isn't open to us and it doesn't ground the narratives of character we tell. The best we can do is to work here from within. At any given point, someone's character is given by the best narrative we can tell about a person that has the appropriate emotional and evaluative significance. This builds in the idea that we can get it wrong and sometimes should revise our judgments. It does not expect us to be omniscient and know all the facts on the ground *a priori*.

Though Goldie's view and my own differ from Aristotle's view of character in several respects, there are some throughlines:

To echo the words of Aristotle in his *Nicomachean Ethics* (1106<sup>b</sup>20), [objectivity in narrative] is a matter of having the right emotions 'at the right times, towards the right people, and in the right way'. The multi-layered multi-level power of narratives to express or reveal perspectives is not a malign power; on the contrary, it is their strength—it is precisely, what we use them for (160).

Because narratives are multi-layered and multi-level, there may not be *the* one single interpretation that holds in a given case. We may even find that there is a range of appropriate reactions and narratives we could tell (which may be relatively narrow or broad depending on the case at hand). However, you can absolutely get narratives of character wrong, and the range of wrong answers is significantly larger than the range of right answers.

---

about others. You can do this both at the level of which details you select and at the level of which interpretation you give of those details. Thanks to Doug Yetman for the initial example.



Finally, character is a representation that we construct, but this does not mean that we change other people's character by simply telling different stories. There are norms internal to the practice that govern what stories we can legitimately tell, and those norms are directly related to what, in fact, is true about the person in question. A person is responsible for their character in that they are the one whose patterns of reason and choice are being evaluated and those patterns delimit accurate narratives of character. Even if that person can't put everything together to see the relevant patterns or make the right interpretation, they are responsible for the details that go into the narrative. Other people may be responsible for choosing to tell and construct narratives of character of others, but it's not as if we directly shifting the facts about others by telling a different story. We are, however, coloring them in a new light that can be assessed and rejected if it isn't fitting. Those narratives and assessments can then influence us in important ways.

## Conclusion

As Talbot Brewer (2009) argues, the values of rich activities may not be immediately accessible to us—we may need to engage in the activity for a long time to slowly discern what makes that activity worthwhile. I suspect that constructing narratives of character is one such rich activity that requires careful attention and that continually draws our attention to new features of moral life.

There is much more to be said about narratives of character than I have had the space to say here. However, I hope that I have laid out at least the basic contours of a view that helps us think more fruitfully about the rich variety of narratives that we can construct and how narratives of character function within moral life. In the final chapter, I turn to questions about some of the limits of responsibility for character on the narrative view.

## Chapter 6: Excuses and Exemptions?

How do excuses and exemptions work on the narrative view? Can we give a clear set of criteria that delineates who is responsible and who isn't?

In this final chapter, I use the case of someone with moral scrupulosity OCD to try to illustrate how my broadly Strawsonian approach informs how we should think about excuses and exemptions. Using Kate Abramson's (2016) view of character as a mode of evaluation, I set the stage for thinking about excuses and exemptions as itself an ethical problem all the way down—one that cannot be solved with a clear taxonomy of psychologically normal and psychologically abnormal agents. I explore David Shoemaker's (2015) project of understanding responsibility at the margins as one of the better attempts to create such a taxonomy and explain why it fails at an impossible task. I then apply Strawson's general categories of the excuses and exemptions to the case of scrupulosity OCD to illustrate how we should implement the general Strawsonian approach. I hope to show that even though I cannot, for good philosophical reason, provide the reader with a taxonomy, there is a way to reasonably determine when the excuses and exemptions hold in individual cases.

### Part 1: Why a Taxonomy Isn't a Good Idea

In this section, I consider whether the moral and medical modes can be neatly lined up so that we get a clean taxonomy of who is responsible in what senses on the basis of certain medical conditions and dis/abilities. I conclude that such a taxonomy is inadvisable given the complex background ethical norms that govern when we should deploy the medical mode and the mode of character.

## 1.1 Character as a Mode

As I argued last chapter, deploying the moral mode and constructing narratives of character is itself a kind of activity done for particular reasons, most notably those with interpersonal import. In this chapter, I follow Kate Abramson's (2016) view that rejects the idea that we can cleanly separate virtues and vices from mental health and illness or natural abilities/inabilities as different natural psychological kinds. Instead, "questions about how to classify a psychological attribute depend on the suitability of various modes of evaluation including moral, medical, and natural ability modes, which differ in content, normative implications, and conditions of application" (56). Ultimately, the "choice of mode of evaluation is itself an ethical choice" (57).

Though I've argued that we do care about the judgments and choices others hold and make as well as the patterns those judgments and choices evince, we tend to select specific sets of judgments and choices for narratives of character because they evoke the reactive attitudes and have a certain kind of interpersonal importance to us. There are internal norms to each mode of evaluation that partially determine the appropriateness of their application:

- To regard someone as having a natural ability is to mark that attribute as a desirable accident of birth, as natural disabilities are undesirable accidents of birth (66).
- To regard a person's attribute as a virtue is to respond to it in a way that marks it as worthy of positive moral attitudes, especially reactive ones: certain forms of praise, love, gratitude. A vice is worthy of (some) negative reactive attitude(s): guilt, shame, anger, resentment, indignation, contempt (66).
- To regard someone as having a mental illness, or as having a healthy psychological attribute, is to say that they are well or ill, could benefit from medical intervention, and the like (66).

More than one mode can be applied in a given situation: my grief during the process of questioning my religious faith could rightly be deemed both a matter of depression and a matter of character—it was a complex process of revisiting my own evaluative commitments and choosing to value my own life despite my strong feelings that I was a worthless sinner. And perhaps the whole process was

grounded in some kind of natural sensitivity and earnestness that made me more susceptible to a deep sense of faith and a deep sense of betrayal when that faith came into question.

A wide variety of our assessments and language combine multiple modes of evaluation. Take those that combine the medical and moral modes, including: “control freak,” “victim complex,” “co-dependent,” “narcissist,” and many more. When appropriate, the medical mode gives us reason to encourage those close to us to seek medical care or therapy or to distance ourselves emotionally from them by taking the objective stance. The moral mode also gives voice to our justified resentments and the wrongs that have been done to us. Sometimes both modes are appropriate; other times they are not. Take the kind of character assessments of laziness or lack of self-control that are wrongly applied to fat and obese people merely because of their weight. Not only is the moral mode uncalled for in this case, but it also frequently interferes with the ability of fat patients to receive medical care. Because everything is seen through the moral lens of “this fat person should lose weight,” other medical issues can be ignored or pushed down the road until the patient is compliant with the request to lose weight.

The question as to whether something should be considered a mental illness or disability “is already, in part, an ethical question” (71). Is trans identity a mental illness because it interferes with the ability to flourish in a gender essentialist society? No. Even the concept of disability itself is grounded in background ethical debates about whether disability should be framed in terms of an individual problem that requires medical intervention and/or a collective and social failure to provide adequate accommodations. Additionally, both the medical mode and the natural ability mode can be used to obscure the ethical and matters of character. Is it just a woman’s hormones? Or is she rightfully angry over being repressed in a patriarchal society?<sup>92</sup> Is the vocal soloist’s ability

---

<sup>92</sup> By contrast, we might imagine a case in which a someone uses a medicalized excuse to avoid any accountability for what is also an ongoing, harmful character trait.

reduceable to pure talent? Or have many years of effort and dedication been devoted to the development of that skill? Even deploying the natural ability mode requires some understanding of the ethical considerations at play (72).

Abramson groups the considerations that help us determine which mode is appropriate into three categories:

- (1) considerations focused on the psychological facts about the agent and the fit between those facts, on the one hand, and the mode of evaluation on the other;
- (2) the interpersonal import of adopting a given mode of evaluation in situ; and
- (3) the comparative appropriateness of adopting one mode of evaluation rather than another in terms of the first two sets of considerations (70).

With regard to (1), certain kinds of basic bodily states do not fit with the mode of character, such as feeling itchy after being assaulted by a number of mosquitos on a hot summer evening. Nor does a fleeting daydream about a flying elephant with a monocle warrant any medical intervention. With regard to (2), the interpersonal import can be quite expansive, including “not only relations between appraiser and appraised, but also others with whom the appraised and appraiser interact, as well as general features of social dynamics perhaps not reducible to interactions between particular persons” (72). This changes from context to context, and different matters of interpersonal import might cut against each other. With regard to (3), though it may sometimes be perfectly appropriate to deploy multiple modes, sometimes one mode is clearly more appropriate than another. The therapist has good reason to deploy a different configuration of modes of evaluation with their client than that client’s close friend would deploy when told the same story from the client’s childhood.

Finally, I should say something about the thin distinction between the appropriateness of deploying the moral mode vs. deploying the mode of character. Since character is more wide-ranging and requires some sort of narrative complexity to get off the ground, it will not be appropriate as a mode of evaluation in cases in which we need only assess whether someone has broken a rule or

whether a single attitude or choice was morally suspect absent any connection to a larger whole.

While the mode of character is a subset of the moral mode, it does not exhaust the moral mode.

I will let Abramson sum up this section:

In all these ways, deciding whether to regard any given attribute as a virtue or vice, a natural ability or defect, a mental illness or a feature of psychological health involves complicated and nested ethical questions all the way down, and at every level. The facts of the agent's psychology on which contemporary ethicists have focused are only one small piece of a complicated puzzle (76).

## 1.2 Carving off the Bounds of Responsibility

There is a common strategy in contemporary accounts of moral responsibility that attempts to clearly show how the view can deal with a range of standard cases where it appears that responsibility is mitigated or removed altogether. David Shoemaker's (2015) book *Responsibility at the Margins* presents perhaps the most nuanced version of this strategy, as he explores a variety of cases of marginal agents in the "perverse sweet spot" where we have ambivalent responsibility responses towards them. Shoemaker proposes that we have roughly three different responsibility concepts that can apply in different degrees to different kinds of agents: attributability responsibility (character), answerability responsibility (judgment), and accountability responsibility (regard).<sup>93</sup> Shoemaker works to determine to what extent individuals with autism, psychopathy, dementia, and other psychological conditions can participate in each of the three kinds of responsibility in Shoemaker's taxonomy.

One of the benefits of Shoemaker's version of this strategy is that it allows for agents with differential ability to participate in one or more of the kinds of responsibility to certain degrees, if not all three. This makes responsibility something that isn't merely on or off but allows for a

---

<sup>93</sup> On Shoemaker's view, each face of responsibility corresponds to a subset of the reactive attitudes: attributability (character) to agential disdain/admiration, answerability (judgment) to agential regret/pride, and accountability (regard) to agential anger/gratitude (35). Each face of responsibility represents a different quality of will. For an earlier version of this view, see Shoemaker (2013).

plurality of notions of moral responsibility. My own view is amenable to multiple complementary conceptions of responsibility, though I have focused primarily on responsibility for character. At the same time, I suspect that even the most nuanced taxonomy that lays out degrees of responsibility and recognizes the diversity of human agency is a strategy that is doomed from the start.

First, it's unclear to me that Shoemaker avoids stigmatizing agents who fail to meet the standards for participating in one or more faces of responsibility. At the best points in Shoemaker's book, like in the discussion of dementia, there is a recognition of the experiences of people with dementia and of their caretakers (210). At the worst points, like in the discussion of psychopathy, it's unclear that Shoemaker avoids making unwarranted moral assumptions about medical conditions. The studies of psychopathy that Shoemaker most thoroughly discusses are of psychopaths who are already in prison, and successful psychopaths are defined as "successful" only in the sense that they are not in prison ... yet" (161).<sup>94</sup> In places, the approach can come off as something like a "philosophical freak show," carving off the bounds of moral agency by looking from the outside at the various inabilities that mark each condition.<sup>95</sup>

Second, Shoemaker's approach, as he considers a range of medically-recognized conditions including clinical depression, mania, OCD, psychopathy, high-functioning autism, mild intellectual disability, and dementia often requires that he take most medical definitions and categories as given. At the broader level, the medical taxonomy itself is often tentative and under development, and it has a number of issues: 1) Taxonomies are tentative and stipulative, even if they are communicated as if they are authoritative and have the final word. 2) Some taxonomies are too restrictive and miss people who have a certain condition because the boundaries have been drawn too narrowly. 3) The

---

<sup>94</sup> The relevant studies are Blair (1995), which used ten incarcerated psychopaths and ten incarcerated non-psychopaths, and Aharoni, Sinnott-Armstrong, and Kiehl (2012), which compared 22 incarcerated psychopaths to 24 incarcerated non-psychopaths.

<sup>95</sup> See August Gorman's (2019, 2020, 2022, 2023) work for better examples of theorizing about responsibility and disability that more thoughtfully prioritize disabled experience.

DSM configures diagnoses symptomatically, which are in turn sometimes defined in terms of cultural norms and expected functioning.<sup>96</sup> Cultural ideas about what counts as an appropriate response to a given situation are themselves often fraught, especially when applied to members of oppressed groups. 4) Current mental health care taxonomies are also large shaped by capitalism and the insurance industry, which may prefer clear, bureaucratic categories over more individually-tailored, rich explanations. The drive towards diagnosing in terms of symptoms also feeds into the needs of insurance companies to have patients diagnosed with conditions that can be “cured” with short-term treatment and lower medical costs.

The sheer range of knowledge that one would have to have to understand the limits of the medical models and the vast range of experiences of individuals who express these conditions differently is near impossible to have for a team of researchers, let alone one author. But the apparent clarity of a taxonomy obscures this vast range of nuanced experience. Even if a taxonomy is helpful as a starting point from which individual moral agents and medical practitioners can deviate in individual cases, it often encourages us to see through that lens and over-simplify cases in terms of the categories we have already been given.<sup>97</sup> I think the particularities of our lives are too rich to be fully captured by such a general frame.<sup>98</sup> As King and May (2018) argue, the variation of symptoms between and within conditions as well as over time within individual agents is too vast to make any generalization from mental illness to responsibility. Mental illnesses are also commonly

---

<sup>96</sup> As Cefalu (2010) notes, “[u]sers of the DSM are tacitly asked to become interpreters of not only individual symptoms but also cultural norms. Every psychiatrist need be something of a sociologist, anthropologist, theologian, and, when extrapolating from stylized conversion narratives and thick descriptions like Bunyan’s, semioticians and literary critics as well” (123). Here, Cefalu is referring to the DSM-IV, specifically in the context of trying to diagnose historical figures with contemporary mental health categories.

<sup>97</sup> See Nguyen (2021a, 2021b, forthcoming) for detailed discussions of bureaucracies, deceptive clarity, and the tradeoffs between creating portable, algorithmic information and exercising individual skill and judgment.

<sup>98</sup> When diagnoses fit and have explanatory power, they can be very helpful for people trying to understand their own experiences and manage their condition. But even this application is often highly personalized and requires a sensitivity to lived experience. For example, there are many varieties of OCD which can be distinguished by their themes, the compulsions involved, and the extent to which the person recognizes what is going on, among other things. Determining which pieces fit and to what extent is highly individualized.



domain specific, and there must be a sufficient connection between the impairment and the action or attitude in question to consider mitigated responsibility (17). King and May conclude that “there is no reason to believe that having a mental disorder generally makes one less responsible than those who enjoy better mental health” (20).

Third, this kind of conditions-internal-to-the-agent framing fails to recognize that disability cannot be understood as entirely internal to the agent.<sup>99</sup> The social model of disability argues that dis/ability is often societally constructed as, for instance, Jay Timothy Dolmage (2018) argues that immigrants at Ellis Island became racialized and disabled when they entered the U.S. and were categorized by officials in ways that selectively allowed certain desirable groups of people in and excluded others.<sup>100</sup> Using the frame of disability is often a choice. In some communities, instead of, for example, medicalizing deafness as something to be cured, the members of that community might instead organize around concepts of Deaf pride and Deaf identity. For all of us, disabled or not, our ability to function (which is itself a complex, value-laden notion) is dependent in part on accommodations, supports, and environment.<sup>101</sup> An amputee who has access to an excellent prosthetic has different modes of functionality than a similar amputee without access to a good prosthetic.

---

<sup>99</sup> Shoemaker (2022) more thoroughly discusses different responsibility-conferring strategies on the basis of the medical and social models.

<sup>100</sup> The medical and social models of disability are usefully contrasting ways of defining disability, though actual views commonly fall somewhere in-between. Per the *SEP* article “Disability: Definitions and Models,” the medical model “explains disability disadvantage in terms of pathological states of the body and mind themselves,” while the social model “explains the characteristic features of disability in terms of a relation between an individual and her social environment: the exclusion of people with certain physical and mental characteristics, or ‘impairment’, from major domains of social life” through the built environment, social structures, and institutions. This relies on a contrast between “*impairments*, individual characteristics or conditions, with *disabilities*, disadvantages imposed by society on those with impairments.”

<sup>101</sup> As Reynolds (2019) writes, “if there is such a thing as *the* ableist fallacy, it is the idea that abilities are individual, that individual bodies ‘have’ abilities. ‘Ability’ is a concept that picks out the relations of body and context, and once one sees it so, the idea of a direct bind between body and worth and the idea of the verdict of bodies is shown to be both unfounded and ethically deplorable” (443).

There is also an exceptionally close connection between disability and oppression. Those in oppressed communities are generally more subject to mental health conditions like depression and ongoing trauma responses because of adverse treatment and hostile environments. And, being disabled opens up more possibilities of becoming further disabled through lack of access to resources or active discrimination. Joel Michael Reynolds (2022) notes that while it is well-known that Black Americans are much more likely to be killed by police, it is less well-known that “nearly half of all people killed by the police are people with disabilities” (48). Reynolds argues that white supremacy also “functions as a process and apparatus of *making* abled and disabled” (48). As this illustrates, questions of when the mode of natural dis/ability should be employed and how social expectations and realities should be considered are already embroiled in ethical considerations.

Fourth, in part because of all of these complex considerations and in part because of my own convictions that good judgments of responsibility require experience and practical judgment, I have abandoned the project of developing a framework that can clearly identify all cases in which people are fully responsible, partially responsible, or not responsible at all. Even if it were possible, I’m unconvinced that such a totalizing project would track what we’re doing when we’re making judgments of responsibility—as I’ve argued in Chapter 5, I think that the project of determining who’s responsible for what is an active, constructive project that we are always engaged in and that we cannot predetermine. As Christine Korsgaard (1996a) argues,

We may think we may be excused from treating certain persons as autonomous and certain governments as legitimate. But in order to decide that there are some human beings to which we will not apply the concepts of full autonomy, or some governments too despotic to count as legitimate, we apparently will have to draw some lines that are not firmly grounded either in theoretical facts about those persons and governments or in the moral law. We must decide who to count as a free rational being, and what makes a state legitimate. These decisions cannot be based on finding out which objects in fact are free rational beings or really embody the general will, because there are no such facts. If applied ethics is done this way, its work is not metaphysical, but ethical and practical all the way down (356).

## Part 2: Strawson and the Limits of Responsibility

In this section, I propose a methodology for applying the Strawsonian excuses and exemptions using moral scrupulosity OCD as a core exemplary case and push back against Strawsonian views that emphasize psychological normality/abnormality as the primary frame for thinking about the excuses and exemptions.

### 2.1 OCD and Responsibility

It's a very complicated thing to read philosophical theories that make you out to be a less than fully responsible agent because you have a mental health condition. While there is a relief in being removed from the obligations of responsibility when it is genuinely unfair to impose them, there is also a very reasonable worry of being cut out of meaningful interpersonal relationships, mutual regard, and love. There is a tendency to think about morally responsible agents in terms of some statistically normal psychological baseline and to then determine how any departures from that baseline impede responsibility. However, I think this is the wrong approach. The core question, I think, is not whether an agent is sufficiently psychologically abnormal or broken to be unable to participate in our practices of moral responsibility but instead whether it is good, appropriate, and helpful to apply the moral mode or other modes in a given case.

In what follows, I use the case of moral scrupulosity OCD to explore how we should think about the excuses and exemptions.<sup>102</sup> However, my general methodological approach has significant upshots that can be applied to agents with other mental health conditions by those who have sufficient knowledge of the particulars. Once we shift our focus from the assumed psychological deficiencies of the agent to the ethical norms that govern which modes we should apply, the

---

<sup>102</sup> The particular case described is my own.

Strawsonian approach becomes much more palatable. This move also seems in line with the general relational and interpersonal focus of Strawson's work.

Imagine someone with scrupulosity OCD who, in times of exceptional stress, experiences obsession and compulsion symptoms for multiple hours every day.<sup>103</sup> This person obsesses over anything regarded as a mistake and frequently feels the compulsion to apologize, confess, seek reassurance, and punish themselves in the hope that they will neutralize the consequences of that mistake and never make it again. This is done out of a fear of catastrophic consequences for even the most minor actions. Frequently, they experience intrusive thoughts of self-harm. Assume that this pattern of self-immolation sometimes interferes with this person's interpersonal relationships. They become so obsessed with determining what exactly they did wrong and what they should do to fix it that they fail to listen to and be truly curious about what is going on in other people's lives. This isn't a constant—it comes and goes depending on other background stresses.

At the same time, this person is involved in their community, has developed deep friendships, and has carefully worked out a variety of coping mechanisms that tend to keep things fairly stable. They understand what is going on, even if they cannot always shift the obsessions and compulsions, and they work carefully to distribute the burden of care so as to not overwhelm any particular relationship. Some of these coping mechanisms themselves almost become part of the obsession and compulsion cycle, as the general pattern of thinking about and reacting to "mistakes" is consistent.

Our expectations of the OCD agent may be reasonably modified depending on the day, the context, and the relationship. However, it may still be appropriate to sometimes hold the agent with scrupulosity OCD to the expectation that they will not constantly seek reassurance from you or

---

<sup>103</sup> Those with OCD often experience common themes around contamination, morality, relationships, and sexuality, among others. There has been a relatively [recent expansion in the diagnostic criteria](#) to include obsessions about real life events and obsessions and compulsions that the agent sees as reasonable.

apologize for every potential slight they might have committed. What governs which expectations are reasonable and when expectations can be placed on the agent? I argue that it is less a matter strictly of the agent's abilities/disabilities and more a matter of what is, all things considered, ethically appropriate and practically feasible.

## 2.2 Excuses

Given the relatively consistent interference in their everyday life that the scrupulous OCD agent experiences, in what way is responsibility mitigated in their case, if at all?<sup>104</sup> If we look at P. F. Strawson's (1962) categories of the excuses and exemptions (localized and global), which categories, if any, apply to the scrupulous OCD agent?<sup>105</sup>

Take the group of excuses, which show that a reactive attitude is inappropriate in a given case without suggesting "that the agent is in any way an inappropriate object of that kind of demand for goodwill or regard which is reflected in our ordinary reactive attitudes." Excuses "do not invite us to see the agent as other than a fully responsible agent. They invite us to see the injury as one for which he was not fully, or at all, responsible." Among the examples Strawson gives for excuses, he lists "He hadn't realized", "He didn't know", as well as "He was pushed", "He had to do it." These excuses generally show that the agent did not harbor the ill will that the action seemed to express.

Take the excuse "He had to do it." I think there is a reasonable case to be made that the OCD agent is operating under a kind of duress. Carl Elliott (1991) suggests that the person with

---

<sup>104</sup> Shoemaker (2015) considers the case of scrupulosity OCD and concludes that scrupulous agents are fully attributable and answerability responsible for many attitudes but mitigated to exempt from responsibility for those that are unconnected to the agent's cares or commitments or those that are not judgment-sensitive. OCD agents may be exempt to mitigated for accountability responsibility depending to the extent to which obsessions block regard (123-124). For another contemporary discussion of responsibility and scrupulosity OCD, see Summers and Sinnott-Armstrong (2015).

<sup>105</sup> Wallace (1994) clarifies the distinction between excuses and exemptions by pointing to A-conditions and B-conditions for responsibility. "B-conditions make it fair to hold people morally to blame for particular things they have done, while A-conditions make it fair to hold people morally accountable. In these terms, excuses indicate the absence of a B-condition of responsibility, while exemptions indicate the absence of an A-condition" (118).

OCD may face a difficult choice between sitting with the anxiety and distress of resisting compulsions and performing a series of irrational actions that the patient may know to be irrational. This mirrors cases of duress, in which the person would strongly prefer to be in a different situation, have different desires, and not face this particular choice. Even if duress does not justify the agent's behavior, it may lessen blame.

Similarly, Steglich-Petersen and Varga (2024) argue that the irrationality of OCD should be understood in terms of harboring recalcitrant emotions. They reject the judgmentalist account of emotion on which, “for example, fearing  $x$  is, in part, a matter of judging or believing  $x$  to be dangerous” (669) and instead adopt Michael Brady's (2009; 2013) version of neojudgmentalism on which emotions are understood “as responses to states of affairs that are of significance to us, and as involving a number of bodily (e.g., visceral, cardiovascular) changes that help mobilize motivational and cognitive resources” (670). For example, “fear constricts the focus of attention on threatening stimuli, renders it demanding to disengage from them, and assigns a lower priority to the processing of neutral information” (670).

On this view,

we should understand emotions as involving inclinations to assent to and act on evaluative construals.... For example, the activation of cognitive resources in anxiety leads to increased attentional focus on signs that support an evaluative construal of the situation as dangerous, inclining the subject to assent to the situation being dangerous and to acting accordingly. Crucially, even if the activation of cognitive and motivational resources inclines a subject in this manner, it is possible for her to resist assenting and acting in ways that the particular construal would dictate. For Brady, this is exactly what characterizes recalcitrant emotional experience: while the subject is inclined to act on and assent to an evaluative construal, she does not act on it and assents instead an opposing construal of her situation (670).

Recalcitrant emotions are thus irrational on Brady's view insofar as they create significant practical and cognitive costs by distracting the agent, requiring the agent to spend more energy to get around the distraction, conflicting with what the agent takes to be true, pushing the agent to assent to

something for which the agent does not think there is good reason to assent, and inviting the agent to invent reasons to justify the emotion (670-1).

For the purposes of my argument, it does not particularly matter whether these recalcitrant “emotions” should be categorized as emotions or something more like urges. It is sufficient to claim that the precipitating psychological factor is not caused by judgment and does not shift even when it seems like it should in response to the agent’s considered judgment. If the distress OCD sufferers experience is not dissimilar to the kind of pressure exacted on coerced people, the obsessions and compulsions then seem to be an understandable response to that kind of pressure, even though it is a pressure with no rational basis.

This makes the language of excuse seem plausibly appropriate, as these details seem to recontextualize the agent’s cares and responses. The particular excuse in this case might read something like “I care deeply about you but I can’t quite seem to shift my focus today. I’m sorry I keep returning to the same thoughts; the OCD is bad today.” Even if blame is only mitigated, this kind of explanation works to show that this action does not communicate a lack of regard for the other person (or at least not as much of a lack of regard as the action suggests).<sup>106</sup>

Whether these excuses mitigate or remove blame entirely depends on other features of the relationship, person, and situation. If the OCD agent is unable to turn away from their obsessive thoughts and compulsions to actively listen even when their partner has just experienced the death of a close family member, the partner may very reasonably blame the OCD agent for failing to make sufficient effort to show and express care. Likewise, if the OCD agent never pays attention to their partner’s needs, it seems very plausible that they do not in fact care deeply for their partner in the

---

<sup>106</sup> Interestingly enough, the kinds of explanations given for relational problems precipitated by conflicts around ADHD, autism, and other mental health and neurodevelopmental conditions often seem to follow the form of excuses rather than exemptions. Chronic lateness in someone with ADHD might be due to time blindness rather than not caring about showing up to class, and a lack of eye contact in an autistic person might simply communicate an aversion to eye contact and not a lack of respect in social situations.

way they assert. If, however, the OCD agent genuinely has a pattern of showing and demonstrating care for their partner regularly and when it matters but cannot seem to do so when they are in a bad spell, it would be unreasonable to blame the OCD agent for failing to be present and demonstrate care all the time. The larger patterns that inform narratives of character frequently modulate our interpretations of and reactions to more localized incidences.

### 2.3 Exemptions

I suspect, however, that the OCD patterns described above more cleanly fit within Strawson's class of exemptions as he describes it. Within this class, Strawson lists two general kinds. As examples of the first kind, Strawson lists: "He wasn't himself" and "He has been under very great strain recently." If exemptions "invite us to suspend our ordinary reactive attitudes towards the agent, either at the time of his action or all the time," then this first kind of exemption appears to be localized. We tend to regard the person in question as generally subject to the reactive attitudes, but he isn't an apt target of the reactive attitudes at the time. For the person who is not currently himself, Strawson notes that: "We normally have to deal with him under normal stresses; so we shall not feel towards him, when he acts as he does under abnormal stresses, as we should have felt towards him had he acted as he did under normal stresses."

For the second kind of exemption, Strawson mentions cases in which the agent is "psychologically abnormal" or "morally undeveloped," potentially including the "neurotic." The examples given for this category include: "That's purely compulsive behavior on his part," "He's only a child," "His mind has been systematically perverted," and "He's a hopeless schizophrenic."<sup>107</sup>

---

<sup>107</sup> Szalai (2016) suggests that OCD agents are not "weak and passive with regard to the compulsive act" but instead that acting on compulsions is voluntary and "regarded by the patient as an instrument of control" (47).



I will call these globalized exemptions, because we suspend our ordinary reactive attitudes all the time. In these cases, “all our reactive attitudes tend to be profoundly modified”:

I must deal here in crude dichotomies and ignore the ever-interesting and ever-illuminating varieties of case. What I want to contrast is the attitude (or range of attitudes) of involvement or participation in a human relationship, on the one hand, and what might be called the objective attitude (or range of attitudes) to another human being, on the other. Even in the same situation, I must add, they are not altogether exclusive of each other; but they are, profoundly, opposed to each other. To adopt the objective attitude to another human being is to see him, perhaps, as an object of social policy; as a subject for what, in a wide range of sense, might be called treatment; as something certainly to be taken account, perhaps precautionary account, of; to be managed or handled or cured or trained; perhaps simply to be avoided, though this gerundive is not peculiar to cases of objectivity of attitude. The objective attitude may be emotionally toned in many ways, but not in all ways: it may include repulsion or fear, it may include pity or even love, though not all kinds of love. But it cannot include the range of reactive feelings and attitudes which belong to involvement or participation with others in inter-personal human relationships; it cannot include resentment, gratitude, forgiveness, anger, or the sort of love which two adults can sometimes be said to feel reciprocally, for each other. If your attitude towards someone is wholly objective, then though you may light him, you cannot quarrel with him, and though you may talk to him, even negotiate with him, you cannot reason with him. You can at most pretend to quarrel, or to reason, with him.<sup>108</sup>

Now, imagine that you're the agent with scrupulosity OCD described above, reading these options for the different kinds of exemptions. It appears that you, as the OCD sufferer, are, in fact, psychologically abnormal and neurotic with compulsive behaviors, which seems to place you in the category of globalized exemptions.

If it turns out that the objective attitude is the only appropriate attitude to adopt towards you, then it seems that, on the Strawsonian view, you cannot be reasoned with and do not have access to ordinary interpersonal human relationships, including those that involve “the sort of love

---

<sup>108</sup> Shoemaker (2022) suggests that we should take the accountability stance to be a subset of the participant stance. This would allow impaired agents to participate in a variety of ordinary interpersonal relationships and be subject to “shared affective experiences of friendliness, rooting, love, amusement, enjoyment, glee, joy, grief, and sadness” even if they are not appropriate targets of accountability (52-53). On Shoemaker’s view, “the participant attitudes are simply about emotional engagement, communion, and connection with other people. They are the glue of humanity” (56). Something about this view seems to be right, insofar as the moral mode does not encompass every aspect of our important relationships with others. Whether it is appropriate to place these attitudes and shared experiences within the participant stance but separate them from accountability or place them somewhere between or outside of the dichotomy between the participant or objective stance, I will leave unanswered for now.

which two adults can sometimes be said to feel reciprocally, for each other.” Yet, this does not seem accurate to the original case as described. If the OCD agent is aware of their condition, they may very well be able to be reasoned with and reject the judgments that their OCD inclines them to. If they are already involved in ordinary interpersonal human relationships and there is no special reason to suspect that those relationships are impaired, why think that they are cut off from the sort of love that two adults can sometimes be said to feel reciprocally for each other?

I suspect that the scrupulosity OCD case I am considering falls somewhere in the category of localized exemptions and that the case is one of the “ever-interesting and ever-illuminating varieties” that Strawson does not have space to consider, one that warrants the objective attitude sometimes and warrants the reactive attitudes at others. Is there a way to read Strawson that can better handle the case of the morally scrupulous OCD agent?

#### 2.4 Ameliorating Strawson?

As I have argued, determining which mode of evaluation is most apt depends on particular psychological facts, other ethical considerations, and matters of relational import, such as what the agent whose responsibility is in question can reasonably demand as a person. This approach seems much more amenable to me than a kind of psychology-first set of categories that delineate who is a full member of the moral community and who is not, based on whether they have normal, developed psychological profiles. Of course, there is a real danger of, on the one hand, overly relying on the medical mode and, on the other, failing to employ the medical mode when it is warranted. If I am right, however, that the scrupulous OCD agent’s experience falls under localized exemptions and that localized exemptions should be seen as a choice of mode rather than a limited exclusion from the moral community, then the Strawsonian view becomes much less threatening.

Contrast this view with Pamela Hieronymi's (2020) reading of Strawson on which "statistically ordinary adult relationships" set the terms for what is reasonable to expect of others in our moral practices (17).<sup>109</sup> Thus, "if we all had lesser natural capacities, things would be different—but not because we would uniformly exempt one another from certain moral demands or responsibility. Instead, the system of demands would, itself, adjust to us—it would adjust to what is typical or tolerably ordinary" (33). Even if Hieronymi's textual interpretation is correct, I think there are good reasons to reject this version of a Strawsonian framework.

First, I think that we can, in fact, adjust our system of demands to the particular needs of individual people. If I turn in a set of revisions late because it was a bad depression week and I was doing my best to manage the condition, no resentment seems warranted—the system of demands should shift to accommodate that. Second, I think it's much easier as an approach to stay largely within our moral practices without having to look for something external to justify or ground them. I took it that, when Strawson says that our practices neither call for nor permit an external "rational" justification, he was getting at something like "if you don't already see the value of human relationships internal to them, you've missed the point."

Whether this is the right way to read Strawson, I don't think a project that cleanly lays out a strict set of capacities or some statistical normality as requirements for full moral membership works at the end of the day. Carving off the abnormal merely because it is statistically abnormal does not seem to have any ethically defensible basis. Likewise, any move to expand the statistically normal to include more palatable kinds of mental health issues leaves out those who are more stigmatized.

Third, by focusing on what we want out of our ethical practices rather than primarily on the

---

<sup>109</sup> Smith and Vickers (2021) similarly argue that "any Strawsonian account of moral responsibility requires full members of the moral community possess the capacity for the interpersonal reactive attitudes, the capacity to see the interpersonal reactive attitudes as demands for a certain kind of treatment or regard, and the capacity to respond to those attitudes" (483). These are the statistically normal capacities they take to be required. Though these capacities seem very plausible to me to be practical preconditions for being able to engage in responsibility exchanges, I am still suspicious of approaches that prioritize capacities over ethical and practical concerns.

dis/abilities of the person in question, we can more clearly see the relational focus of morality rather than trying to spell out some idealized, normal moral agent from which all other non-responsible agents deviate.

How do we apply these insights to the OCD agent? When they are in a time of increased stress, the OCD symptoms also tend to increase. In these instances in which the OCD patterns get the best of them, the objective attitude of management may be the right mode to adopt. When they are stuck in an anxiety loop, it may not do much good to talk at the level of responsibility for their apparent judgments of what is fear-worthy because that's not necessarily what's going on underneath—they may reject their worries to no avail. It also wouldn't do for the OCD agent to blame themselves for an increase in obsessions if they are taking all reasonable steps to manage them and improve their condition. If anything, an attitude like self-gratitude might be warranted by the attempts to heal.

When the OCD agent is in calmer times, there may be little reason to withhold the normal reactive attitudes or merely see them as an object of treatment. When the anxieties are not activated, it may be helpful for the OCD agent to reason at the level of responsibility to try to reassess, reframe, and push back against some of their obsessions and compulsions. Even if the OCD agent will always find it very difficult to deal with resentment directed towards them, that does not mean that resentment is always inappropriate. When the OCD agent takes up excessive swaths of time and energy from others while seeking reassurance, they may very well be an apt target of resentment. When that resentment is communicated, there is still an opportunity for the OCD agent to take it seriously and work to change their behavior in order to improve the relationship. Though OCD often makes ordinary interpersonal relationships harder, it does not preclude them altogether.

## Conclusion

The case of an agent with moral scrupulosity OCD is meant to illustrate how my broader methodology can be applied to a particular case. I have explicitly avoided including agents with other impairments, as it would take more particular knowledge of each to do them justice. Similarly, my observations about this case of moral scrupulosity OCD do not necessarily hold for all cases of moral scrupulosity OCD, let alone all cases of OCD. Once, however, we appreciate how the Strawsonian reactive attitudes can be applied or withheld on the basis of immediate features of a situation and the larger patterns that they are a part of, we find ourselves with a flexible and responsive system of moral responsibility that does justice to the particular people involved and that starts from an ethical perspective. This system is necessarily informed by and responsive to experience, and it works best when applied to specific contexts rather than to general categories of ability or disability. You cannot bypass these ethical questions and determine responsibility on the basis of ability alone.

## Conclusion

There are a number of methodological insights my view proposes that should inform not only how we think about responsibility for character but also how we should think about moral responsibility in general.

First, moral responsibility is something we have to work out on the ground with particular people and situations. There is no top-down, god's-eye-view perspective from which we can solve everything. There is no strict set of abilities that we can identify that are perfectly lined up with who is responsible in a given sense and who is not. And whatever criteria for responsibility we adopt, they need to be something that may reasonably enter into our everyday ascriptions of responsibility.

Second, because holding others responsible is something that is motivated by particular practical concerns, we are not necessarily morally responsible for every choice or judgment or pattern of behavior. Only those that warrant gratitude, resentment, or those reactive attitudes that are around and between enter into the realm of moral responsibility.

Third, character is not a distinct face of responsibility that should be wholly separated from moral accountability for judgment-sensitive attitudes and choices. Though character requires some sort of narrative complexity that goes beyond a single attitude or choice, it is not fundamentally in a different realm of responsibility. Instead, it is a whole that is made up in part of other states that we normally take it that people can be responsible for. Given that we care quite a bit about the general patterns that people exhibit in our interpersonal relationships, character has a central role to play in our ordinary practices of moral responsibility. It is not secondary to some more basic kind of responsibility.

Fourth, while choice and control matter for narratives of character in a way that cannot be reduced to judgment, we need to reject any fetishization of choice, force of will, or complete individual sovereignty. We are not isolated, perfectly controlled agents but actual human beings who

are affected by the people and systems around us and who are often internally divided in sometimes productive ways. As much as our unreflective attitudes and judgments can cause us grief, sometimes our inner impulses push us towards the good. We need both executive control and our inner wellsprings of sensitivity to balance each other. We do not need to theorize perfect, ideally responsible agents to get a compelling picture of moral responsibility, and, in my view, the more we allow for the mess inside and around us, the better.

Fifth, adding in more context and particulars do not necessarily mitigate responsibility. Understanding immediate historical context, environment, relationships, and other features of someone's life can equally provoke empathy and antipathy, depending on the details. These rich narratives can lead to more nuanced conclusions, or they may hit us over the head with the blunt realization that someone is a total asshole.

Sixth, the fluid quality of narratives of character is a feature and not a bug. We construct narratives of character for immediate practical reasons, and those narratives can be accurate or inaccurate, charitable or uncharitable at any given point in time. There are clear norms that govern whether those narratives of character are accurate and charitable, and those narratives frequently must be updated as the details change or as our legitimate reasons to tell them change.

There are a number of questions I have not had the space to touch on in this project:

- Which interpretive language can legitimately be applied to narratives of character, and what are the conditions for application?
- Which narratives of character are in fact accurate and charitable?
- What other particular cases help us get a sense of the wider range of excuses and exemptions?
- What do my conclusions suggest about responsibility and free will?
- How many different kinds of responsibility are there?

- What conclusions might hold for legal contexts of thinking about responsibility?
- How do narratives of character relate to questions about personal identity?
- What might my view of responsibility imply for non-human species that seem to share at least some core features of our moral responsibility practices?

Though these questions remain unanswered for now, I hope that I have provided a compelling way to approach questions of moral responsibility for character that places human judgment and experience at the center.



## References

- Abramson, Kate. 2016. "Character as a Mode of Evaluation." *Oxford Studies in Normative Ethics* 6: 56-76. <https://doi.org/10.1093/acprof:oso/9780198790587.003.0004>.
- Abramson, Kate, and Leite, Adam. 2011. "Love as a Reactive Emotion." *Philosophical Quarterly* 61 (245): 673-699. <https://doi.org/10.1111/j.1467-9213.2011.716.x>.
- Adams, Robert. 1985. "Involuntary Sins." *The Philosophical Review* 94 (1): 3-31. <https://doi.org/10.2307/2184713>.
- Aharoni, Eyal, Sinnott-Armstrong, Walter, and Kiehl, Kent A. 2012. "Can Psychopathic Offenders Discern Moral Wrongs? A New Look at the Moral/Conventional Distinction." *Journal of Abnormal Psychology* 121: 484-97.
- Annas, Julia. 2005. "Comments on John Doris's Lack of Character." *Philosophy and Phenomenological Research* 71 (3): 636-642.
- . 2011. *Intelligent Virtue*. Oxford University Press.
- Arpaly, Nomy. 2003. *Unprincipled Virtue: An Inquiry into Moral Agency*. Oxford University Press.
- Arpaly, Nomy, and Schroeder, Timothy. 2013. *In Praise of Desire*. Oxford University Press.
- . 1999. "Praise, Blame and the Whole Self." *Philosophical Studies* 93 (2): 161-188. <https://doi.org/10.1023/a:1004222928272>.
- Baron, Marcia. 2018. "Shame and Shamelessness." *Philosophia* 46: 721-731. <https://doi.org/10.1007/s11406-017-9933-x>.
- Betzler, Monika. 2009. "Expressive Actions." *Inquiry* 52 (3): 272-292. <https://doi.org/10.1080/00201740902917143>.
- Blair, R. J. R. 1995. "A Cognitive Developmental Approach to Morality: Investigating the Psychopath." *Cognition* 57: 1-29.
- Brady, Michael S. 2009. "The Irrationality of Recalcitrant Emotions." *Philosophical Studies* 145: 413-430.
- . 2013. *Emotional Insight: The Epistemic Role of Emotional Experience*. Oxford University Press.
- Brewer, Talbot. 2009. *The Retrieval of Ethics*. Oxford University Press.
- . 2021. "Acknowledging Others." *Journal of Ethical Reflections* 1 (4): 91-119.
- Cefalu, Paul. 2010. "The Doubting Disease: Religious Scrupulosity and Obsessive-Compulsive Disorder." *Journal of Medical Humanities* 31: 111-125.

- Cherry, Myisha. 2021. *The Case for Rage: Why Anger is Essential to Anti-Racist Struggle*. Oxford University Press.
- ContraPoints. 2021. "J. K. Rowling." YouTube, January 26, 2021.  
[https://www.youtube.com/watch?v=7gDKbT\\_l2us](https://www.youtube.com/watch?v=7gDKbT_l2us).
- Currie, Gregory. 2010. *Narratives and Narrators: A Philosophy of Stories*. Oxford University Press.
- Darwall, Stephen. 2006. *The Second-Person Standpoint: Morality, Respect, and Accountability*. Harvard University Press.
- . 2015. "Second-Personal Attitudes of the Heart (By Featured Philosopher, Steve Darwall)." PEA Soup. Published April 6, 2015.  
<https://peasoup.typepad.com/peasoup/2015/04/second-personal-attitudes-of-the-heart-by-featured-philosopher-steve-darwall.html>.
- Deonna, Julien A. 2006. "Emotion, Perception, and Perspective." *Dialectica* 60 (1): 29-46.  
<https://doi.org/10.1111/j.1746-8361.2005.01031.x>.
- Dockterman, Eliana. 2022. "You're Wrong About Host Sarah Marshall Was Always Fascinated by History's Misunderstood Women. Now the Rest of Us Are Too." *Time*, February 22, 2022.  
<https://time.com/6149560/youre-wrong-about-sarah-marshall-interview/>.
- Dolmage, Jay Timothy. 2018. *Disabled Upon Arrival: Eugenics, Immigration, and the Construction of Race and Disability*. The Ohio State University Press.
- Eliot, George. (1879) 2023. *Theophrastus Such*. Createspace.
- Elliott, Carl. 1991. "Moral Responsibility, Psychiatric Disorders, and Duress." *Journal of Applied Philosophy* 8 (1): 45-56.
- Fischer, John Martin, and Ravizza, Mark. 1998. *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge University Press.
- Fischer, John Martin, Tognazzini, Neal A. 2009. "The Truth about Tracing." *Noûs* 43 (3): 531-556.  
<https://doi.org/10.1111/j.1468-0068.2009.00717.x>.
- Frankfurt, Harry. 1971. "Freedom of the Will and the Concept of a Person." *The Journal of Philosophy* 68 (1): 5-20. <https://doi.org/10.2307/2024717>.
- . 1992. "The Faintest Passion." *Proceedings and Addresses of the American Philosophical Association* 66 (3): 5-16.
- . 2004. *The Reasons of Love*. Princeton University Press.
- Goldie, Peter. 2012. *The Mess Inside: Narrative, Emotion, and the Mind*. Oxford: Oxford University Press.

- Gorman, August. 2019. "The Minimal Approval View of Attributability." In *Oxford Studies in Agency and Responsibility 6*, edited by David Shoemaker. Oxford University Press.
- . 2020. "Depression's Threat to Self-Governance." *Social Theory and Practice* 46 (2): 277-297.
- . 2022. "Demystifying the Deep Self View." *Journal of Moral Philosophy* 19 (4): 390-414.
- . 2023. "What is the Difference between Weakness of Will and Compulsion?" *Journal of the American Philosophical Association* 9 (1): 37-52.
- . n.d. "Book." Accessed November 21, 2023. <https://www.augustgorman.com/book>.
- Greco, Lorenzo. 2015. "The Self as Narrative in Hume." *Journal of the History of Philosophy* 53 (4): 699-722. <https://doi.org/10.1353/hph.2015.0086>.
- Hieronymi, Pamela. 2019. "I'll Bet You Think This Blame is About You." In *Oxford Studies in Agency and Responsibility Volume 5: Themes From the Philosophy of Gary Watson*, edited by Justin Coates and Neal Tognazzini, 60-87. Oxford University Press.
- . 2020. *Freedom, Resentment, and the Metaphysics of Morals*. Princeton University Press.
- Hochschild, Arlie Russell. 2012. *The Managed Heart*. University of California Press.
- Hume, David. 1978. *A Treatise of Human Nature, 2nd Edition*. Edited by Lewis A. Selby-Bigge and Peter H. Niddich. Oxford: Clarendon Press.
- . 1998. *An Enquiry Concerning the Principles of Morals*. Edited by Tom L. Beauchamp. Oxford: Oxford University Press.
- . 1999. *An Enquiry Concerning Human Understanding*. Edited by Tom L. Beauchamp. Oxford: Oxford University Press.
- Kamtekar, Rachana. 2004. "Situationism and Virtue Ethics on the Content of Our Character." *Ethics* 114 (3): 458-491. <https://doi.org/10.1086/381696>.
- King, Matt. 2020. "Attending to Blame." *Philosophical Studies* 177 (5): 1423-1439. <https://doi.org/10.1007/s11098-019-01260-w>.
- King, Matt, and May, Joshua. 2018. "Moral Responsibility and Mental Illness: A Call for Nuance." *Neuroethics* 11 (1): 11-22. <https://doi.org/10.1007/s12152-017-9345-4>.
- Korsgaard, Christine M. 1996a. *Creating the Kingdom of Ends*. Cambridge University Press.
- . 1996b. *The Sources of Normativity*. Cambridge University Press.
- Locke, Don, and Frankfurt, Harry. 1975. "Three Concepts of Free Action." *Aristotelian Society Supplementary Volume* 49 (1): 95-126.

- MacIntyre, Alasdair. (1981) 2007. *After Virtue: Third Edition*. University of Notre Dame Press.
- Manne, Kate. 2017. *Down Girl*. Oxford University Press.
- Mason, Michelle. (2003). "Contempt as a Moral Attitude." *Ethics* 113 (2): 234-272.  
<https://doi.org/10.1086/342860>.
- . (2006). "Aretaic Appraisal and Practical Reason." *Southern Journal of Philosophy* 44: 629-656.
- . (2010). "On Shamelessness." *Philosophical Papers* 39 (3): 401-425.  
<http://dx.doi.org/10.1080/05568641.2010.538916>.
- McKenna, Michael. 2012. *Conversation and Responsibility*. Oxford University Press.
- Moody-Adams, Michele. 1994. "Culture, Responsibility, and Affected Ignorance." *Ethics* 104 (2): 291-309.
- Nagel, Thomas. 1979. *Mortal Questions*. Cambridge University Press.
- Nguyen, C. Thi. 2021a. "The Seductions of Clarity." *Royal Institute of Philosophy Supplement* 89: 227-255. <https://doi.org/10.1017/S1358246121000035>.
- . 2021b. "Transparency is Surveillance." *Philosophy and Phenomenological Research* 105 (2): 331-361. <https://doi-org.proxyiub.uits.iu.edu/10.1111/phpr.12823>.
- . (forthcoming). "Value Capture." *Journal of Ethics and Social Philosophy*.
- Prinz, Jesse. 2004. *Gut Reactions: A Perceptual Theory of Emotion* Oxford and New York: Oxford University Press.
- Reynolds, Joel Michael. 2019. "The Meaning of Ability and Disability." *Journal of Speculative Philosophy* 33 (3): 434-447.
- . 2022. "Disability and White Supremacy." *Critical Philosophy of Race* 10 (1): 48-70.
- Robert, P.H., et al. 2002. "The Apathy Inventory: Assessment of Apathy and Awareness in Alzheimer's Disease, Parkinson's Disease, and Mild Cognitive Impairment." *International Journal of Geriatric Psychiatry* 17: 1099-105.
- Romano, Aja. 2023. "Is J.K. Rowling Transphobic? Let's Let Her Speak for Herself." *Vox*, March 16, 2023. <https://www.vox.com/culture/23622610/jk-rowling-transphobic-statements-timeline-history-controversy>.
- Scanlon, T. M. 1998. *What We Owe to Each Other*. Cambridge: Harvard University Press.
- . 2008. "Blame." In *Moral Dimensions*, 122-214. Harvard University Press.
- Schechtman, Marya. 1996. *The Constitution of Selves*. Cornell University Press.

- Schroeder, Mark. 2007. *Slaves of the Passions*. Oxford University Press.
- Schroeder, Timothy, and Arpaly, Nomy. 1999. "Alienation and Externality." *Canadian Journal of Philosophy* 29 (3): 371-387. <https://doi.org/10.1080/00455091.1999.10717517>.
- Shoemaker, David. 2011. "Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility." *Ethics* 121: 602-632. <https://doi.org/10.1086/659003>.
- . 2013. "Qualities of Will." *Social Philosophy and Policy* 30 (1-2): 95-120. <https://doi.org/10.1017/S0265052513000058>.
- . 2015. *Responsibility from the Margins*. Oxford: Oxford University Press.
- . 2022. "Disordered, Disabled, Disregarded, Dismissed: The Moral Costs of Exemptions from Accountability." In *Agency in Mental Disorder: Philosophical Dimensions*. Edited by Matt King and Joshua May. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198868811.003.0003>.
- Schulman, Sarah. 2016. *Conflict is Not Abuse: Overstating Harm, Community Responsibility, and the Duty of Repair*. Arsenal Pulp Press.
- Smart, J. J. C. 1961. "Free-will, Praise, and Blame." *Mind* 70 (279): 291-306.
- Smith, Angela. 2005. "Responsibility for Attitudes: Activity and Passivity in Mental Life." *Ethics* 115 (2): 236-271. <https://doi.org/10.1086/426957>.
- . 2012. "Attributability, Answerability, and Accountability: In Defense of a Unified Account." *Ethics* 122 (3): 575-589. <https://doi.org/10.1086/664752>.
- . 2008. "Control, Responsibility, and Moral Assessment." *Philosophical Studies* 136: 367-392.
- . 2015a. "Responsibility as Answerability." *Inquiry* 58 (2): 99-126. <https://doi.org/10.1080/0020174X.2015.986851>.
- . 2015b. "Attitudes, Tracing, and Control." *Journal of Applied Philosophy* 32 (2): 115-132. <https://doi.org/10.1111/japp.12107>.
- Smith, Nicholas, and Vickers, Darby. 2021. "Statistically Responsible Artificial Intelligences." *Ethics and Information Technology* 23: 483-493. <https://doi.org/10.1007/s10676-021-09591-1>.
- Spelman, Elizabeth. 1978. "On Treating Persons as Persons." *Ethics* 88 (2): 150-161.
- Steglich-Petersen, Asbjørn, and Varga, Somogy. (2024). "Obsessive-Compulsive Disorder and Recalcitrant Emotion: Relocating the Seat of Irrationality." *Philosophical Psychology* 37 (3): 658-683.
- Strawson, Galen. 1994. "The Impossibility of Moral Responsibility." *Philosophical Studies* 75: 5-24.

- Strawson, P. F. 1962. "Freedom and Resentment." *Proceedings of the British Academy* 48: 1-25.
- Substance Abuse and Mental Health Services Administration (US). 2016. "Impact of the DSM-IV to DSM-5 Changes on the National Survey on Drug Use and Health," Table 3.13, DSM-IV to DSM-5 Obsessive-Compulsive Disorder Comparison. Rockville (MD), June 2016. <https://www.ncbi.nlm.nih.gov/books/NBK519704/table/ch3.t13/>
- Summers, Jesse S., and Sinnott-Armstrong, Walter. 2015. "Scrupulous Agents." *Philosophical Psychology* 28 (7): 947–966. <http://dx.doi.org/10.1080/09515089.2014.949005>.
- Szalai, Judit. 2016. "Agency and Mental States in Obsessive-Compulsive Disorder." *Philosophy, Psychiatry, & Psychology* 23 (1): 47-59. <https://doi.org/10.1353/ppp.2016.0001>.
- Tessman, Lisa. 2005. *Burdened Virtues: Virtue Ethics for Liberatory Struggles*. Oxford University Press.
- Vargas, Manuel. 2005. "The Trouble with Tracing." *Midwest Studies in Philosophy* 29 (1): 269-291. <https://doi.org/10.1111/j.1475-4975.2005.00117.x>.
- Wallace, R. Jay. 1994. *Responsibility and the Moral Sentiments*. Cambridge: Harvard University Press.
- . 2019. *The Moral Nexus*. Princeton University Press.
- Wasserman, David, and Sean, Aas. 2022 (2011). "Disability: Definitions and Models." *The Stanford Encyclopedia of Philosophy (Fall 2023 Edition)*. Edited by Edward N. Zalta and Uri Nodelman. <https://plato.stanford.edu/archives/fall2023/entries/disability/>.
- Watson, Gary. (1987) 1993. "Responsibility and the Limits of Evil: Variations on a Strawsonian Theme." In *Perspectives on Moral Responsibility*, edited by John Martin Fischer and Mark Ravizza, 119-148. Cornell University Press.
- . (1996) 2004. "Two Faces of Responsibility." In *Agency and Answerability*, 227-248. Oxford: Oxford University Press.
- Wedgwood, Ralph. 2015. "The Pitfalls of 'Reasons'." *Philosophical Issues* 25 (1): 123-143. <https://doi.org/10.1111/phis.12054>.
- Williams, Bernard. 1981. *Moral Luck*. Cambridge University Press.
- . 1985. *Ethics and the Limits of Philosophy*. Cambridge: Harvard University Press.
- Wolf, Susan. (1987) 2003. "Sanity and the Metaphysics of Responsibility." In *Free Will*, edited by Gary Watson, 372-387. New York: Oxford University Press.
- . 1990. *Freedom Within Reason*. Oxford University Press.
- . 2015a. "Responsibility, Moral, and Otherwise." *Inquiry* 58 (2): 127-142.

———. 2015b. “Character and Responsibility.” *The Journal of Philosophy* 112 (7): 356-372.

Zheng, Robin. 2016. “Attributability, Accountability, and Implicit Bias.” In *Implicit Bias and Philosophy, Volume 2: Moral Responsibility, Structural Injustice, and Ethics*, 63-89. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198766179.003.0004>.

## Curriculum Vitae

Elizabeth Cargile Williams

### EDUCATION

---

<b>Indiana University</b> – Bloomington, IN Ph.D. Philosophy	2017 – 2024
<b>University of Tennessee</b> – Knoxville, TN M.A. Philosophy	2015 – 2017
<b>University of Tennessee</b> – Knoxville, TN B.A. Music	2010 – 2014

---

### EMPLOYMENT

---

- Co-Founder** – philosophy for humans (May 2023-Present)  
Organized a public-facing free Zoom conference on AI ethics which drew over 250 sign-ups from academics and non-academics, keynote given by philosopher C. Thi Nguyen
- News Analyst** – *The Prindle Post*, The Prindle Institute (January 2022-December 2022)  
Wrote articles on ethics in the news tailored towards Ethics Bowl participants
- Ethics Bowl Coach** – Indiana University, Bloomington (2021-2022)  
1<sup>st</sup> Place, Upper Midwest Regional Competition, Fall 2021
- Associate Instructor** – Indiana University, Bloomington  
P242 Applied Ethics, Summer 2022 – Primary Instructor: Elizabeth Williams  
P141 Intro to Ethical Theories, Spring 2022 – Lecturer: John Robison  
P242 Applied Ethics, Fall 2021 – Primary Instructor: Elizabeth Williams  
P140 Intro to Ethics: Bioethics, Spring 2020 – Primary Instructor: Elizabeth Williams  
P393 Bioethics, Fall 2019 – Lecturer: John Robison  
P140 Intro to Ethics: Bioethics, Spring 2019 – Professor: Dr. Sandra Shapshay  
Philosophy 100, Fall 2018 – Professor: Dr. David McCarty
- Graduate Teaching Assistant** – University of Tennessee, Knoxville  
Philosophy 101, Spring 2017 – Professor: Dr. Adam Cureton  
Philosophy 101, Fall 2016 – Lecturer: Dr. Joshua L. Watson  
Philosophy 101, Spring 2016 – Professor: Dr. Adam Cureton  
Contemporary Moral Problems, Fall 2015 – Professor: Dr. Adam Cureton
-



## DISSERTATION PROJECT

---

### **Dissertation Title** – “Character as Narrative: Moral Responsibility in Context”

Advisor: Dr. Kate Abramson

*Abstract:* In this dissertation, I propose a theory of moral responsibility for character, understood as a kind of narrative. Narratives of character give evaluative and emotional coherence to the judgments and choices of others understood within context. In Chapter 1, I set the stage for the project by arguing that Gary Watson’s (1996) notion of accountability, properly understood, requires some reference to the particular person, rather than merely roles or rules. In Chapter 2, I push back against R. Jay Wallace’s (1994) view of responsibility that makes us out to be responsible only for our choices or the things we can directly control. In Chapter 3, I argue that T. M. Scanlon’s (1998) view rightly captures responsibility for judgment-sensitive attitudes but wrongly underplays the distinct importance of choice and control. In Chapter 4, I consider Susan Wolf’s (1990) and Fischer and Ravizza’s (1998) historical requirements on responsibility and conclude that the former too quickly removes agency and responsibility from those who have chosen to go down evil paths while the latter isn’t something we have access to in our everyday moral practices of accountability. In Chapter 5, I draw on Peter Goldie’s (2012) work on narratives to create a novel view of character, understood not as a stable disposition but as an actively constructed narrative built from a number of elements we care about, including our choices, our judgments, our history and context, as well as other features of what we’re like. These narratives often draw out patterns that have interpersonal import, and they can range from very small character studies to narratives that span a person’s life. Finally, in Chapter 6, I argue that constructing a taxonomy that clearly delineates who is responsible from who is not is a project that is doomed from the start, and I consider what my broadly Strawsonian approach can tell us about the excuses and exemptions.

---

## MASTERS THESIS

---

### **Thesis Title** – “Rational Engagement as a Way of Showing Respect to Oneself and Others”

Advisor: Dr. Adam Cureton

*Abstract:* We often encounter people with seemingly unreasonable theoretical beliefs: for example, persons who are global warming deniers, creationists, and conspiracy theorists. I explore how respect informs our response to these persons. I conclude that we ought to at least be willing or disposed to engage in rational discussion sometimes with even seemingly unreasonable persons as a way of recognizing and respecting their rational nature. I take up the underlying philosophical issues of how theoretical reason relates to respect, how considerations of self-respect should be taken into account in non-ideal circumstances, and how respect can require us to act in everyday interpersonal interactions.

---

## PUBLICATIONS

---

**Feminist Philosophical Quarterly** (2022) – “Why You Ought to Defer: Moral Deference and Marginalized Experience”

Co-Author: Savannah Pearlman  
Volume 8, No. 2

---

## PRESENTATIONS

---

**American Philosophical Association** – January 4-7, 2023

Conference Title: Eastern Division Meeting

Presentation Title for Accepted Colloquium Paper: “Separating Moral Accountability from Substantive Responsibility”

*Abstract:* In this paper, I provide an argument that suggests we should separate accountability from substantive responsibility within the domain of moral responsibility. On my view, moral accountability requires a specific kind of interpersonal address that is lacking in substantive responsibility. These two kinds of responsibility may overlap, but they are distinct. Whereas substantive responsibility primarily concerns the principles that govern the fair distribution of burdens and obligations, moral accountability is something that happens between individual persons and carries with it an expressive function. Substantive responsibility does not require any assessment of character or practical identity; moral accountability does. As such, moral accountability requires attributability. This revised notion of moral accountability points us in a virtue theoretic direction, re-centering character and re-emphasizing the kinds of interpersonal interactions at the core of our moral practices.

**Stanford Ethics Workshop** – 2020-2021

Paper Accepted: “Why You Ought to Defer: Epistemic Deference, Actional Deference, and Marginalized Identity”

Co-Author: Savannah Pearlman

Did not present due to the date being moved back significantly due to COVID/paper already being submitted for publication.

**Athena in Action** – June 17-19, 2020

Presentation Title: “Why You Ought to Defer: Epistemic Deference, Actional Deference, and Marginalized Identity”

Co-Presenter: Savannah Pearlman

**American Philosophical Association** – April 17-20, 2019

Conference Title: Pacific Division Meeting

Presentation Title for Invited Panel: “Identity-Based Moral Harms”

Co-Presenter: Savannah Pearlman

*Abstract:* In this paper, we make the case for deferring to an individual's testimony of their experience with identity-based moral harms. By deferring, we mean accepting an individual's testimony as a credible source of justification and conforming one's actions in accordance with the testifier's recommendation(s). By identity-based moral harms, we mean those moral harms which individuals experience qua their identity as a member of a marginalized group. We argue that these individuals are moral experts with regard to the identity-based moral harms that they have themselves experienced. We contend that failure to defer to individuals in these instances demonstrates a blatant disregard for the testifier as a moral and epistemic agent. On this basis, we will provide a response to the understanding objection to moral deference, which states that one should not act on moral prescriptions without understanding why one should act in such a way.

**OUTstanding Seminar** – October 22, 2016

Conference Title: Let it all OUT

Presentation Title: "Epistemic Injustice and the Imperative of (LGBT+) Inclusion in Higher Education"

Co-Presenter: Alex Richardson

**Society for Ethnomusicology, Southeast and Caribbean Chapter** – March 13-14, 2015

Conference Title: Peaks and Valleys

Paper Title: "A Tale of Two Churches: Forming and Expressing Appalachian and Gender Identities Through Church Music"

Advisor: Dr. Rachel Golden

---

COURSES TAUGHT AS PRIMARY INSTRUCTOR

---

**Applied Ethics: Ethics Bowl**

Indiana University, Bloomington

Fall 2021, Summer 2022 (Taught as Intensive Writing Class)

*Course Description:* Through critically analyzing a variety of case studies, students will develop their ethical reasoning skills and learn to engage in respectful, productive ethical discourse about controversial topics that span a wide variety of disciplines. Students will be provided methodological tools for analysis that will help with identifying the morally relevant features of a case, asking good questions that get at the core moral issues that the case raises, brainstorming creative solutions to problems, applying theoretical moral considerations to a case, and formulating arguments to answer pressing ethical questions. Students will also practice participating in collective ethical inquiry and discussion and will learn how to productively engage in conversations about difficult ethical topics.

The skills and values taught in this course are in line with those promoted by the Intercollegiate Ethics Bowl. A team of students will be constructed from this class (and from others participating in the tryouts) to compete in the Regional and National IEB competitions. Participation in the Ethics Bowl is neither required nor guaranteed for

students in this class; prospective team members must express interest and try out for the team.

### **Intro to Ethics: Bioethics, Life at the Margins**

Indiana University, Bloomington  
Spring 2020

*Course Description:* In this course we will discuss ethical issues that arise at the beginning and end of life: Is abortion morally permissible? Should physicians assist patients in ending their lives? How do we deal with cases in which children and the elderly cannot give consent? In triage situations, is it permissible to discriminate based upon age? To help us understand these questions, we will employ three central ethical theories: virtue ethics, utilitarianism, and deontology. Students will survey a variety of issues in bioethics and consider difficult, real-life cases encountered by those working in the medical field. Students will learn how to think carefully and critically about complex ethical issues and discover how some of our best ethical theories would resolve them.

---

### HONORS AND AWARDS

---

#### **Clark Outstanding AI Award – May 2022**

Description: Annual award given to one graduate instructor in the philosophy department.

#### **Moss Fellowship – August 2020 to May 2021**

Description: Awarded a dissertation fellowship given to two graduate students in the philosophy department.

#### **Graduate Academic Excellence Award – May 2021**

Description: Annual award given to one graduate student in the philosophy department.

#### **Ford Foundation Fellowship Honorable Mention**

Description: Did not receive a fellowship, but was awarded an honorable mention for a meritorious application.

#### **Irving and Shirley Brand Graduate Fellowship – August 2018 to May 2019**

Description: Awarded fellowship given to one graduate student in the Humanities with a preference for students in philosophy.

#### **1<sup>st</sup> and 2<sup>nd</sup> Year Clark Essay Prize Winner – Spring 2018**

Description: Awarded monetary prize for best essay among 1<sup>st</sup> and 2<sup>nd</sup> year graduate students in philosophy.

#### **First-Year Fellowship Recipient – August 2017 to May 2018**

Description: Awarded one of two first-year fellowships granted to incoming philosophy graduate students at Indiana University, Bloomington.

**Haslam Scholars Program** – June 2010 to May 2014

Description: Premier academic scholarship program at the University of Tennessee. Selected as one of fifteen high school students to receive a full scholarship for academic achievement.

**Phi Beta Kappa** – April 2013 to May 2014

Description: Academic honors society for the liberal arts and sciences.

**Omicron Delta Kappa** – August 2011 to May 2014

Description: National leadership honors society.

President 2013-2014

Vice President 2012-2013

**Honors Council** – August 2010 to May 2014

Description: Governing student body for the Chancellor's Honors Program at the University of Tennessee.

Secretary 2011-2013

Freshman Representative 2010-2011

**Pi Kappa Lambda** – April 2014 to May 2014

Description: National music honors society

---

PROFESSIONAL SERVICE

---

**Blog of the APA** – Contributed article “Rage Against the Machine,” on February 9<sup>th</sup>, 2023

**Admissions Representative** – Philosophy Department, Indiana University, Bloomington, Spring 2022

**Co-Social Chair** – Philosophy Department, Indiana University, Bloomington, Fall 2021 to Spring 2022

**NHSEB Judge** – Nationals, Preliminary Round 2, Parr Center for Ethics, Spring 2021

**Social Chair** – Philosophy Department, Indiana University, Bloomington, Fall 2019 to Spring 2020

**Climate Committee Member** – Philosophy Department, Indiana University, Bloomington, Fall 2018 to Spring 2019

**Session Chair** – Indiana Philosophical Association, Fall 2017 Meeting

**Social Chair** – Philosophy Department, University of Tennessee, Knoxville, August 2016 to August 2017

---